

Single-Call Mechanisms

Christopher A. Wilkens*

Balasubramanian Sivan†

Abstract

Truthfulness is fragile and demanding. It is oftentimes computationally harder than solving the original problem. Even worse, truthfulness can be utterly destroyed by small uncertainties in a mechanism’s outcome. One obstacle is that *truthful payments depend on outcomes other than the one realized*, such as the lengths of non-shortest-paths in a shortest-path auction. Single-call mechanisms are a powerful tool that circumvents this obstacle — they implicitly charge truthful payments, guaranteeing truthfulness in expectation using only the outcome realized by the mechanism. The cost of such truthfulness is a trade-off between the expected quality of the outcome and the risk of large payments.

We largely settle when and to what extent single-call mechanisms are possible. The first single-call construction was discovered by Babaioff, Kleinberg, and Slivkins [BKS10] in single-parameter domains. They give a transformation that turns any monotone, single-parameter allocation rule into a truthful-in-expectation single-call mechanism. Our first result is a natural complement to [BKS10]: we give a new transformation that produces a single-call VCG mechanism from any allocation rule for which VCG payments are truthful. Second, in both the single-parameter and VCG settings, we precisely characterize the possible transformations, showing that a wide variety of transformations are possible but that all take a very simple form. Finally, we study the inherent trade-off between the expected quality of the outcome and the risk of large payments. We show that our construction and that of [BKS10] simultaneously optimize a variety of metrics in their respective domains.

Our study is motivated by settings where uncertainty in a mechanism renders other known techniques untruthful. As an example, we analyze pay-per-click advertising auctions, where the truthfulness of the standard VCG-based auction is easily broken when the auctioneer’s estimated click-through-rates are imprecise.

1 Introduction

In their seminal work that sparked the field of Algorithmic Mechanism Design, Nisan and Ronen [NR01] made a striking observation: naïvely computing VCG payments for shortest-path auctions requires computing “ n versions of the original problem.” In their case, it requires solving $n + 1$ different shortest path problems in a network. Over the next decade, as researchers studied computation in mechanisms, they repeatedly noticed that computing payments is harder than solving the original problem. Babaioff et al. [BBNS08] exhibited a problem for which deterministic truthfulness is precisely $(n + 1)$ -times harder than the original problem. In the case of Nisan and Ronen’s own path auction, Hershberger et al. [HSB07] showed that computing VCG prices for a directed graph requires time equivalent to \sqrt{n} shortest path computations.¹

*Computer Science Dept., University of California at Berkeley. cwilkens@cs.berkeley.edu. Supported in part by NSF grant CC-0964033 and by a Google University Research Award. Part of this work was done while the author was an intern at Microsoft Research, Redmond, WA.

†Computer Sciences Dept., University of Wisconsin - Madison. balu2901@cs.wisc.edu. Supported in part by NSF award CCF-0830494. Part of this work was done while the author was an intern at Microsoft Research, Redmond, WA.

¹ Interestingly, the undirected case is easier. Hershberger and Suri [HS01, HS02] show that it only requires time equivalent to a single shortest-path computation. Their work is orthogonal to our own — single-call mechanisms achieve truthfulness in a

Surprisingly, Babaioff, Kleinberg, and Slivkins [BKS10] recently showed that randomization eliminated this difficulty for a large class of problems. They showed that, if in a single-parameter domain payments need only be *truthful in expectation*, then they may be computed by solving the original problem only once. They apply their result to Nisan and Ronen’s path auctions to get a truthful-in-expectation mechanism that uses precisely one shortest-path computation and chooses the shortest path with probability arbitrarily close to 1. We call this a *single-call mechanism*.

The usefulness of Babaioff, Kleinberg, and Slivkins’ result goes far beyond speeding up computation: *Their construction enables truthfulness in cases in which computing “ n versions of the original problem” is informationally impossible.* To use again the Nisan-Ronen path auction, suppose that the graph represents a packet network with existing traffic. In this case, the actual transit times (i.e. costs to edges) may be increased by congestion. While it is possible to estimate congestion ex ante, it is generally impossible to precisely know its effect without transmitting a packet and explicitly measuring its transit time. Unfortunately, since VCG prices depend on the transit times for many different paths, naively computing them will inherit any estimation errors. Even worse, when bidders have conflicting beliefs about such errors, *naively computing “VCG” prices with bad estimates may not guarantee truthfulness* even if the errors are small enough that they not affect the path chosen by the mechanism. In such a case, truthfulness may be regained using a mechanism that only requires measurements along a single path, that is, a mechanism that only requires measurements returned by a single call to the shortest-path algorithm. We will concretely demonstrate this phenomenon later using an example based on pay-per-click advertising auctions.

An important question arises then: *In which mechanism design problems, and to what extent, are single-call mechanisms possible?* In this paper we study, and largely settle, this question. First, we show that this it is possible to transform any mechanism that charges VCG prices in expectation into a roughly equivalent single-call mechanism. While similar in spirit to [BKS10], our reduction charges prices that are fundamentally different from the mechanism in that paper — they do not coincide even when applied to the same allocation rule. Second, we give characterization theorems, delineating precisely the single-call mechanisms that are possible, for both the VCG and single-parameter settings. Finally, single-call constructions offer a tradeoff between expectation and risk. Our characterization theorems allow us to derive lower bounds on this tradeoff, establishing that our VCG construction and the construction of [BKS10] are optimal in a general sense.

Mechanisms, Allocations, and Payments One cornerstone of mechanism design is the decomposition of a mechanism into two distinct parts: an allocation function and a payment function. This approach has borne much fruit — it first revealed fundamental relationships between allocation functions and their nearly unique truthful prices, and it subsequently allowed researchers to study the the two problems in isolation. Like [BKS10], we leverage this decomposition to study payment techniques that apply to large classes of allocation functions — naturally, our primary requirement is that the allocation function may only be evaluated once.

We will focus on single-call mechanisms for two classes of allocation functions that, together, comprise most allocation functions for which truthful payments are known: *monotone single-parameter functions* and *maximal in distributional range (MIDR) functions*.

An allocation function is said to be *single-parameter* if an agent’s bid can be expressed as a single number. This setting was first studied by Myerson [Mye81] in the context of single-item auctions. Subsequent generalizations showed that truthful prices existed if and only if a single-parameter allocation is monotone and provided an explicit characterization of truthful payments. We will use one such characterization

limited-information setting using only one shortest-path computation, while [HSB07, HS01, HS02] assume complete information and study an algorithmic problem.

developed by Archer and Tardos [AT01].

An allocation function is said to be *maximal in distributional range* (MIDR) if, for some fixed set of distributions over outcomes, the allocation always chooses one that maximizes the social welfare of the bidders. MIDR allocation functions are important because they are precisely the ones for which VCG payments are truthful [DD09].

Truthfulness Under Uncertainty Our motivation for developing and optimizing single-call mechanisms comes from scenarios where nature prohibits computing an allocation more than once, most often due to parameter uncertainty. We give a few examples here; more generally, we conjecture that most mechanism design problems have similar variants.

In the uncertain shortest-path auction described earlier, truthful prices will depend on the incremental effect of transit times adjusted for congestion. If the auctioneer generates the network traffic, he may be able to predict the congestion in an edge better than the edge itself and use this prediction when computing the shortest path. However, each edge may individually disagree with the auctioneer’s estimate, and these beliefs are generally unknown to the auctioneer. If the auctioneer were to simply compute VCG payments by combining his estimates with players’ bids, the prices would likely not be truthful. On the other hand, we can require that payments are computed using measured transit times instead of estimates; however, it is informationally impossible to know the precise delay along edges that were not actually traversed. A single-call mechanism sidesteps this hurdle by using only the delays along traversed edges for which the delay had been precisely known.

Machine scheduling offers another application for single-call mechanisms. In some applications (e.g. cloud services), it is common for machines to bid in terms of cost per unit time (or other resource). It is then the responsibility of the scheduler to estimate the time required for the job on that machine. If the scheduler’s estimates differ from a machine’s belief about a job’s runtime, then we find ourselves in the same situation as the path auction — the standard truthful prices for this single-parameter setting will depend on machines’ beliefs about the runtimes of jobs under alternate schedules. A single-call mechanism sidesteps this problem because it requires only the runtimes of jobs under the schedule chosen by the mechanism, which may be measured.

Another interesting example arises in the application of learning procedures such as multi-arm-bandits (MABs). In recurring mechanisms, it is natural for the auctioneer to run a learning algorithm across multiple auctions. For example, when an online advertising auction is repeated, the auctioneer tries to learn the likelihood that a particular ad will get clicked. Computing truthful prices requires knowing what would have happened if the learner had been initialized with a different set of bids. This setting was the original motivation of [BKS10], where they showed that their single-call construction allowed a MAB to be implemented truthfully with $O(\sqrt{T})$ regret. This contrasts with results of Babaioff, Sharma, and Slivkins [BSS09] and Devanur and Kakade [DK09] who showed that any universally truthful mechanism must have regret at least $\Omega(T^{\frac{2}{3}})$ for different measurements of regret.

Finally, in Section 5 we analyze *single-shot pay-per-click (PPC) advertising auctions*. A PPC advertising auction ranks bidders using their pay-per-click bid (i.e. they only pay when they receive a click) and an estimate of the probability of a click (the click-through rate, or CTR). If the bidders’ estimates of their own CTRs are different from the auctioneer’s, truthful prices necessarily depend on bidders’ beliefs about the CTRs, which are unknown.

Single-Call Mechanisms and Reductions Our tool for creating single-call mechanisms is *the single-call reduction*, the main object of study in this paper. A single-call reduction is a transformation that takes an allocation function as a black box and produces a truthful-in-expectation mechanism that calls the allocation

function once. Since the expected payment is equal to the truthful payment for the resulting mechanism, the payments are dubbed *implicit*.

Babaioff, Kleinberg, and Slivkins [BKS10] discovered such a reduction for single-parameter domains. Using only the guarantee that the black-box allocation rule is monotone, their reduction produces a truthful-in-expectation mechanism that implements the same outcome as the original allocation rule with probability arbitrarily close to 1.²

VCG is a mechanism design framework much broader than single-parameter. *Can we construct similar single-call mechanisms that charge VCG prices?* We answer this in the affirmative by giving a reduction producing, for any MIDR allocation function, a single-call mechanism that charges VCG prices in expectation. Analogous to [BKS10], our reduction transforms any MIDR allocation rule into a truthful-in-expectation mechanism that implements the same outcome as the original allocation rule with probability arbitrarily close to 1. However, our construction is fundamentally different in that the distribution of payments does not coincide with [BKS10] when an allocation is both MIDR and single-parameter. This reduction can guarantee truthfulness in multi-parameter mechanisms with uncertainty, as described above, and can also be used to speed up payment computation in MIDR settings like Dughmi and Roughgarden’s [DR10] truthful FPTAS for welfare-maximization packing problems.

We next ask *what single-call reductions are possible?* Babaioff et al. generalize to a class of self-resampling procedures. Subsequent research [Har11] generalized further (and simplified substantially), but concisely characterizing single-call reductions remained an open question. We give tight characterization theorems, showing that a wide variety of reductions are possible and that payments have a very simple characterization in both scenarios. The key technical idea is a simple proof equating a reduction’s expected payments with those required for truthfulness, giving a sharp characterization of the parameters in the reduction. Our technique is a very simple alternative to the contraction mapping argument in [BKS10].

Finally, we ask *what are the best single-call reductions?* As noted above, known single-call reductions choose an outcome different from the original allocation rule with some small probability δ . The penalty for making δ small is that the payments may occasionally be very large — we study this tradeoff. Our study is not unprecedented: [BKS10] asked, as an open question, if their reduction optimized payments with respect to the welfare loss, and Lahaie [Lah10] show a similar tradeoff between the size and complexity of kernel-based payments achieving ϵ -incentive compatibility in single-call combinatorial auctions.

We study the tradeoff inherent to single-call mechanisms with respect to three measures of expectation — welfare, revenue, and a technical (but natural) precision metric — and two measures of risk — variance and worst-case payments. We show that our VCG reduction and the single-parameter reduction of [BKS10] *simultaneously* optimize the tradeoff between expectation and risk for all these criteria.

2 Preliminaries

A mechanism is a protocol among n rational agents that implements a social choice function over a set of outcomes \mathcal{O} . Agent i has preferences over outcomes $o \in \mathcal{O}$ given by a *valuation function* $v_i : \mathcal{O} \rightarrow \mathbf{R}$. The function v_i is private but is drawn from a publicly known set $V_i \subseteq \mathbf{R}^{\mathcal{O}}$.

A *deterministic direct revelation mechanism* \mathcal{M} is a social choice function $A : V_1 \times \dots \times V_n \rightarrow \mathcal{O}$, also known as an *allocation rule*, and a vector of payment functions P_1, \dots, P_n where $P_i : V_1 \times \dots \times V_n \rightarrow \mathbf{R}$ is the amount that agent i pays to the mechanism designer. When a direct revelation³ mechanism is instantiated,

²The authors of [BKS10] have observed that their construction may be extended to any domain where the bid space is convex.

³“Direct revelation” means that an agent’s bid b_i is an element of V_i . In general this need not be the case; however, by the revelation principle, any social choice rule that may be truthfully implemented may be implemented as a direct revelation mechanism that charges the same payments in equilibrium.

each agent reports a *bid* $b_i \in V_i$. The mechanism uses bids $b = (b_1, \dots, b_n)$ to choose an outcome $A(b) \in \mathcal{O}$ and to compute payments $P_i(b)$. The utility $u_i(v_i, o)$ that agent i receives is $u_i(v_i, o) = v_i(o) - P_i$. A mechanism is *truthful* (or incentive compatible) if bidding truthfully (i.e. $b_i = v_i$) is a dominant strategy. Formally, for each i , each $v_{-i} \in V_{-i}$, and every $v_i, v_i' \in V_i$, we have $u_i(v_i, A(v)) \geq u_i(v_i', A(v_{-i}, v_i'))$, where v_{-i} denotes the vector of valuations for all agents except agent i .

A mechanism is *ex-post individually rational* (IR) if agents always get non-negative utility, and mechanism has *no positive transfers* (NPT) if for each agent i and each $v \in V$, $P_i(v) \geq 0$, i.e., the mechanism never pays a player money.

A randomized mechanism is a distribution over deterministic mechanisms. Thus, $A(b)$ and $P_i(b)$ are random variables. For randomized mechanisms, properties like truthfulness may be said to hold universally or in expectation. A randomized mechanism is *universally truthful* if it is truthful for every deterministic mechanism in its support. It is *truthful in expectation* if, in expectation over the randomization of the mechanism, truthful bidding is a dominant strategy. Henceforth, we use truthful, IR, and NPT to mean truthful in expectation unless otherwise noted.

MIDR Allocation Rules MIDR mechanisms are variants of *VCG mechanisms*, mechanisms that maximize social welfare and charge “VCG payments”. Formally, a VCG mechanism’s social choice rule satisfies $A(v) \in \operatorname{argmax}_{o \in \mathcal{O}} \sum_j v_j(o)$, and its payments are $P_i(v) = h_i(v_{-i}) - \sum_{j \neq i} v_j(A(v))$ for some function $h_i : V_{-i} \rightarrow \mathbb{R}$. VCG payments are the only universal technique known to induce truthful bidding. The most common implementation of VCG payments uses the Clarke-Pivot payment rule: set $h_i(v_{-i}) = \max_{o \in \mathcal{O}} (\sum_{j \neq i} v_j(o))$, which gives the only payments that simultaneously satisfy truthfulness, IR, and NPT.

More generally, any allocation rule that maximizes an affine function of agents’ valuations can be truthfully implemented with VCG payments. Moreover, Roberts’ theorem [Rob79] implies that in a general setting (when $V_i = \mathbf{R}^{\mathcal{O}}$), if A is onto (every outcome can be realized), then A has truthful payments if and only if it is an affine maximizer. If the “onto” restriction is relaxed, a social choice function is truthfully implementable with VCG payments if and only if it is (weighted) maximal-in-range (MIR) [NR07] or, for randomized mechanisms, maximal-in-distributional-range (MIDR) [DD09]:

Definition 1 *An allocation rule A is MIDR if there is a set \mathcal{D} of probability distributions over outcomes such that A outputs a random sample from the distribution $D \in \mathcal{D}$ that maximizes expected welfare. Formally, for each $v \in V$, $A(v) = o \sim D^*$ where $D^* \in \operatorname{argmax}_{D \in \mathcal{D}} \mathbf{E}_{o \sim D} [\sum_i v_i(o)]$.*

A weighted MIDR allocation rule maximizes the weighted social welfare $\sum_i w_i v_i(o)$ for $w_i \geq 0$.

Single-Parameter Domains A larger class of social choice rules can be implemented when V_i is single dimensional. We say that a social choice rule has a single-parameter domain if $v_i(o) = t_i f_i(o)$ for some publicly known function $f_i : \mathcal{O} \rightarrow \mathbb{R}_+$. The value $t_i \in T_i$ is an agent’s type (T_i is her type-space, and $T = T_1 \times \dots \times T_n$), and submitting i ’s bid precisely requires stating $b_i = t_i$. When $T = \mathbb{R}_+^n$, we say that bidders have *positive types*. We also use $A_i(b) = f_i(A(b))$ as shorthand, and we say A is bounded if the functions A_i are bounded functions.

A single-parameter social choice rule may be implemented if and only if it is *monotone*, where $A : T \rightarrow \mathcal{O}$ is said to be monotone if for each agent i , for all $b_{-i} \in T_{-i}$ and for every two bids $b_i \geq b_i'$, we have $A_i(b_i, b_{-i}) \geq A_i(b_i', b_{-i})$. This was first shown for a single item auction by Myerson [Mye81]; Archer and Tardos [AT01] gave the current generalization:

Theorem 2.1 [Myerson + Archer-Tardos] *For a single parameter domain, an allocation rule A has truthful payments (P_1, \dots, P_n) if and only if A is monotone. These payments take the form*

$$P_i(b) = h_i(b_{-i}) + b_i A_i(b_i, b_{-i}) - \int_0^{b_i} A_i(u, b_{-i}) du,$$

where $h_i(b_{-i})$ is independent of b_i .

These payments simultaneously satisfy IR and NPT if and only if $P_i^0(b_{-i}) = 0$. Such a mechanism is said to be normalized.

3 Single-call mechanisms

We call a mechanism a single-call mechanism if it only evaluates the allocation function once:

Definition 2 *A single-call mechanism \mathcal{M} for an allocation rule A is a truthful mechanism that has only oracle access to A and computes both the allocation and payments with a single call to A .*

To construct a single-call mechanism, we must first specify the possible allocation functions A and then construct one procedure that yields a single-call mechanism for any A in this set. Thus, the tool for creating a single-call mechanism is a single-call reduction:

Definition 3 *A single-call reduction is a procedure that takes any allocation function A from a fixed set (as a black box) and returns a single-call mechanism.*

For example, the procedure of [BKS10] is a single-call reduction that takes any A drawn from the set of all monotone, bounded, single-parameter allocation rules and returns a single-call mechanism. Similarly, our construction for VCG prices is a single-call reduction that takes any A that is MIDR and returns a single-call mechanism.

To formalize single-call reductions, we first note the following requirements:

- A reduction must take a bid vector b and a black-box allocation function A as input.
- A reduction must evaluate A on at most one bid vector \hat{b} , causing the outcome $A(\hat{b})$ to be realized.⁴
- A reduction must charge payments λ_i that are a function of b , \hat{b} , and $A(\hat{b})$ (and possibly its own randomness).

These requirements suggest the following generic definition of a single-call reduction to turn an allocation function A into a truthful-in-expectation single-call mechanism $\mathcal{M} = (\mathcal{A}, \{\mathcal{P}_i\})$:

1. Solicit the bid vector b from agents.
2. Use b to compute the modified bid vector \hat{b} . This implicitly defines a probability measure $\mu_b(B)$ denoting the probability of choosing $\hat{b} \in B \subseteq V_1 \times \dots \times V_n$ as the modified (resampled) bid vector when b is the actual bid vector. When $\hat{b}_i \neq b_i$, we say that i 's bid was resampled.

⁴Strictly speaking, there may be settings where a single-call reduction could realize an outcome other than $A(\hat{b})$. However, our restriction follows naturally in scenarios where “computing $A(b)$ ” means realizing $A(b)$ and making measurements. It is also required for complete generality because there is no reason to believe that the designer knows how to realize any outcome other than $A(\hat{b})$.

ALGORITHM 1: Generic Single-Call Reduction $(\mu, \{\lambda_i\})$

input : Black box access to an allocation function A , which is drawn from a known set.

output: Truthful-in-expectation mechanism $\mathcal{M} = (\mathcal{A}, \{\mathcal{P}_i\})$.

- 1 Solicit bid vector b from agents;
 - 2 Sample $\hat{b} \sim \mu_b$;
 - 3 Realize the outcome $A(\hat{b})$; // $\mathcal{A}(b)$ is the random function $A(\hat{b})$ where $\hat{b} \sim \mu_b$
 - 4 Charge payments $\lambda(A(\hat{b}), \hat{b}, b)$; // $\mathcal{P}_i(b)$ is the random function $\lambda_i(A(\hat{b}), \hat{b}, b)$ where $\hat{b} \sim \mu_b$
-
3. Declare the outcome to be $A(\hat{b})$, i.e. evaluate A at the modified bid vector \hat{b} . This implicitly defines the allocation function $\mathcal{A}(b)$ which samples $\hat{b} \sim \mu_b$ and chooses the outcome $A(\hat{b})$. The resampling procedure must ensure that truthful payments $\mathcal{P}(b)$ exist for $\mathcal{A}(b)$; Note that $\mathcal{A}(b)$ and $\mathcal{P}(b)$ are random variables that depend on the randomly resampled bid vector \hat{b} . Also, $\mathcal{A}(b)$ and $\mathcal{P}(b)$ are randomized even if $A(b)$ and $P(b)$ are deterministic;
 4. Use b , \hat{b} , and $A(\hat{b})$ to compute payments $\lambda_i(A(\hat{b}), \hat{b}, b)$ that satisfy truthfulness in expectation, that is, charge player i a payment $\lambda_i(A(\hat{b}), \hat{b}, b)$ such that $\mathbf{E}_{\hat{b}}[\lambda_i(A(\hat{b}), \hat{b}, b)] = \mathbf{E}_{\hat{b}}[\mathcal{P}_i(b)]$.

This general procedure is illustrated in Algorithm 1.

We describe a single-call reduction in the above framework by the tuple $(\mu, \{\lambda_i\})$, where μ implies specifying the resampling measure μ_b for all $b \in V_1 \times \dots \times V_n$. Since payments should be finite, we require that λ_i be finite everywhere, and we also require that it be integrable. For the rest of this paper, we assume that λ_i 's are deterministic. For randomized λ_i 's, the characterization theorems still hold with λ_i 's replaced by their expectations over the randomness used.

We say that a reduction is *normalized* if $b_i(A(b)) = 0$ for all i implies $\lambda_i(A(\hat{b}), \hat{b}, b) = 0$, i.e. when every agent receives zero value, all payments are zero.

3.1 Optimal Reductions — Expectation vs. Risk

There are two downsides to the mechanisms produced by single-call reductions. First, there is a penalty in *expectation*, i.e., the expected outcome $\mathbf{E}_{\hat{b}}[A(\hat{b})]$ produced by the reduction is not identical to the desired outcome, $A(b)$. This modified outcome may reduce the expected welfare or revenue of the mechanism, or it may simply cause it to do the “wrong” thing.

Second, there is a penalty in *risk* because the payments λ may vary significantly, i.e. for a fixed b the payments at different resampled bids \hat{b} could be very different. In particular, the magnitude of the payment charged by the single-call mechanism may be much larger than the payments in the original mechanism, i.e. it may be that $|\lambda_i| \gg |P_i|$ for certain outcomes.

Our characterization theorems reveal that there is a fundamental trade-off between expectation and risk. Thus, we call a reduction optimal if it minimizes risk with respect to a lower bound on the expectation.

3.1.1 Expectation

We study three criteria for measuring the expectation of a reduction: $\Pr(\hat{b} = b|b)$, social welfare, and revenue.

The first criterion, $\Pr(\hat{b} = b|b)$ (the precision), measures the likelihood that the reduction modifies players' bids. This criterion is natural when modifying bids is inherently undesirable:

Definition 4 The precision of a reduction α_P is the probability that the reduction does not alter any player's bid:

$$\alpha_P \equiv \min_b \Pr(\hat{b} = b|b) .$$

The other criteria measure standard quantities in mechanism design:

Definition 5 The welfare approximation α_W of a single-call reduction is given by the worst-case ratio between the welfare of the single-call mechanism and the welfare of the original allocation function:

$$\alpha_W = \min_{A,b} \frac{\mathbf{E}_{\hat{b}} [\sum_i b_i(\mathcal{A}_i(b))]}{\sum_i b_i(A_i(b))} .$$

When the welfare of A is zero, $\alpha_W = 1$ if the welfare of \mathcal{A} is also zero and unbounded otherwise.

Definition 6 The revenue approximation α_R of a single-call reduction is given by the worst-case ratio between the revenue of the single-call mechanism and the revenue of the original allocation function:

$$\alpha_R = \min_{A,b} \frac{\mathbf{E}_{\hat{b}} [\sum_i P_i(b)]}{\sum_i P_i(b)} .$$

When the revenue of A is zero, then $\alpha_R = 1$ when the revenue of \mathcal{A} is also zero and unbounded otherwise.

In the case of continuous spaces we replace min/max with inf/sup as appropriate for infinite domains.

3.1.2 Risk

We measure risk through both the variance of payments and their worst-case magnitude.⁵ In order to make a meaningful comparison across different allocation functions and bids, we normalize by players' bids:⁶

Definition 7 Decompose λ_i into terms which depend only on the payoff to a single bidder j (i.e. on $b_j(A(\hat{b}))$ instead of $A(\hat{b})$):

$$\lambda_i(A(\hat{b}), \hat{b}, b) = \sum_j \lambda_{ij}(b_j(A(\hat{b})), \hat{b}, b)$$

(our characterizations in Sections 4 and 6 show that this is possible for our settings). Then the bid-normalized payments of the reduction are given by

$$\sum_j \frac{\lambda_{ij}(b_j(A(\hat{b})), \hat{b}, b)}{b_j(A(\hat{b}))} .$$

We can thus write the variance of bid-normalized payments as

$$\max_{A,i} \text{Var}_{\hat{b} \sim \mu_b} \left(\sum_j \frac{\lambda_{ij}(b_j(A(\hat{b})), \hat{b}, b)}{b_j(A(\hat{b}))} \right)$$

and the worst-case magnitude as

$$\max_{A,i,\hat{b}} \left| \sum_j \frac{\lambda_{ij}(b_j(A(\hat{b})), \hat{b}, b)}{b_j(A(\hat{b}))} \right|$$

where we replace min/max with inf/sup as appropriate for infinite domains.

⁵Intuition suggests optimizing with respect to a high-probability bound. Unfortunately, this is problematic because ignoring low-probability events can dramatically change the expected payment. Thus, in general it is not reasonable to conclude a priori that low-probability events can be ignored.

⁶Intuition also suggests normalizing by the truthful prices for A (i.e. by P_i), but constant allocation functions such as $A_i(b) = 1$ have $P_i = 0$, making this impossible. Bid-normalized payments are a next logical choice.

3.1.3 Optimality

We define an optimal reduction as one that *simultaneously* optimizes the six-way trade-off between expectation and risk:

Definition 8 *A single-call reduction optimizes the variance of/worst-case payments with respect to precision/welfare/revenue for a set of allocation functions if for every bid b , it minimizes the variance of/worst-case normalized payments over all possible reductions that achieve a precision of α_P / welfare approximation of α_W / revenue approximation of α_R .*

4 Maximal-in-distributional-range reductions

In this section, we show how to construct a single-call reduction for MIDR allocation rules, i.e. we show how to construct a randomized, truthful mechanism from an arbitrary MIDR allocation rule A using only a single black-box call to A . The main results are Theorem 4.1, a characterization of all reductions that use VCG payments for an arbitrary MIDR allocation rule, and an explicit construction that optimizes the expectation-risk tradeoff.

Truthful payments for MIDR allocation rules are given by VCG payments with the Clarke-Pivot rule:⁷

$$\mathbf{E}[p_i] = \mathbf{E}[\text{total welfare of bidders without } i] - \mathbf{E}[\text{total welfare of bidders } j \neq i \text{ with } i] \quad (1)$$

(where the expectation is over the randomization in the given MIDR allocation rule). The reduction comes from this formula for $\mathbf{E}[p_i]$: we need to measure the welfare without agent i (the first term in the RHS), so, with some probability, we ignore agent i and maximize the welfare of the remaining agents. Intuitively, this is equivalent to evaluating the allocation function where i 's bid is changed to a “zero” bid while other bids remain the same.

Unfortunately, having removed agent i , even with a small probability, means that computing truthful payments for agent $j \neq i$ requires knowing the allocation where both i and j are ignored. By induction, a single-call mechanism must generate all sets of agents $M \subseteq [n]$ with some probability. Thus, we get an intuitive picture of the reduction's behavior: it will randomly pick a set of bidders $M \subseteq [n]$ and zero the bids of agents not in M .

4.1 Characterizing Truthfulness

We consider reductions in which i 's resampled bid \hat{b}_i is always b_i or zero,⁸ where “zero” means that the agent has a valuation of zero for all outcomes. That is, the resampling measure $\mu_b(B)$ represents a discrete distribution over the bids $\{\hat{b}^M\}$ where $M \subseteq [n]$ is a set of agents and

$$\hat{b}_i^M = \begin{cases} b_i & i \in M \\ 0 & i \notin M \end{cases}$$

Resampling to \hat{b}^M is equivalent to ignoring the welfare of agents outside M and evaluating A at b .

⁷If we relax the no positive transfers requirement, a trivial way to construct a single-call mechanism is to ignore the first term in (1). However, the resulting mechanism would make a huge loss because no agent would ever pay the mechanism.

⁸Even if explicit “zero” bids are not known to the reduction, we assume that the reduction can induce A to optimize the utility of an arbitrary subset of agents. Note that a black-box allocation function can only be turned into a truthful mechanism (even if multiple calls to A are allowed) if it can ignore at least one bidder at a time, so our assumption is not unreasonable.

In the most general setting, our restriction to zeroing reductions is without loss of generality because b and zero are the only bids that are guaranteed to be valid inputs to A for all MIDR allocation functions A . That said, even if a multi-parameter bid structure were known, VCG payments do not depend on the outcome at any other bid. Thus, intuition suggests that resampling to other bids will not be helpful even if it is possible. This intuition can be formalized, but we do not do it here.

Let $\pi(M)$ be a distribution over sets $M \subseteq [n]$. We define the *associated coefficients* $c_i^\pi(M)$ as:

$$c_i^\pi(M) = \begin{cases} -1, & i \in M \\ \frac{\pi(M \cup \{i\})}{\pi(M)}, & i \notin M \end{cases}$$

Intuitively, c_i^π is the weighting that ensures $-\pi(M \cup \{i\})c_i^\pi(M \cup \{i\}) = \pi(M)c_i^\pi(M)$ (where $i \notin M$) to match the terms in (1).

We prove the following characterization of all truthful MIDR reductions $(\pi, \{\lambda_i\})$ that work for all MIDR A :

Theorem 4.1 *A normalized single-call reduction, with VCG payments, for the set of all MIDR allocation rules satisfies truthfulness, individual rationality, and no positive transfers in an ex-post sense if and only if it takes the form $(\pi, \{\lambda_i\})$ where $\pi(M)$ is a distribution over sets $M \subseteq [n]$, the coefficients $c_i^{\pi(M)}$ are finite, and payments take the form*

$$\lambda_i(A(\hat{b}^M), \hat{b}^M, b) = c_i^\pi(M) \sum_{j \neq i} b_j(A(\hat{b}^M)) .$$

Proof: Recall that in general, a multi-parameter allocation function that can be rendered truthful by VCG payments must be MIDR. Thus, our reduction must ensure that \mathcal{A} is MIDR, and we first derive the implications of this requirement on the single-call reduction. We have already assumed that $\mu_b(B)$ is a distribution over bids $\{\hat{b}^M\}$. Let $\pi_b(M)$ be the probability of selecting \hat{b}^M given b .

First, we show that \mathcal{A} is always MIDR if and only if $\pi_b(M)$ does not depend on b . For the *if* direction, if $\pi_b(M)$ is independent of b then \mathcal{A} is a distribution over MIDR allocation rules, and by [DR10], such an allocation rule is MIDR.

For the *only if* direction, we use contradiction. Assume that there are some bids x and y such that $\pi_x(M) \neq \pi_y(M)$ for some M . Then there exists a set $S \subseteq [n]$ such that $\Pr_{\pi}(M \subseteq S|x) \neq \Pr_{\pi}(M \subseteq S|y)$ (by contradiction and induction, start with $S = \emptyset$). Consider an allocation function that has welfare $\sum_i b_i(A(\hat{b}^M)) = 0$ for $M \subseteq S$ and $\sum_i b_i(A(\hat{b}^M)) = 1$ otherwise. The welfare of \mathcal{A} will be precisely $1 - \Pr_{\pi}(M \subseteq S)$, implying that for either x or y , \mathcal{A} did not chose the distribution that maximized social welfare and is therefore not MIDR. Thus, the allocation rule \mathcal{A} is MIDR for all MIDR A if and only if $\mu_b(B)$ is a discrete distribution $\pi(M)$ independent of b .

Next, we write VCG payments for \mathcal{A} that satisfy individual rationality and no positive transfers using the Clarke-Pivot payment rule:

$$\begin{aligned} \mathbf{E}[\mathcal{P}_i] &= \sum_{j \neq i} \sum_{M \subseteq [n]} \pi(M) b_j(A(\hat{b}^{M \setminus \{i\}})) - \sum_{j \neq i} \sum_{M \subseteq [n]} \pi(M) b_j(A(\hat{b}^M)) \\ &= \sum_{M | i \notin M} \pi(M \cup \{i\}) \sum_{j \neq i} b_j(A(\hat{b}^M)) - \sum_{M | i \in M} \pi(M) \sum_{j \neq i} b_j(A(\hat{b}^M)) . \end{aligned} \quad (2)$$

By definition of $\lambda_i(A(\hat{b}^M), \hat{b}^M, b)$, we know that the expected payment made by i will be

$$\mathbf{E}[\mathcal{P}_i] = \sum_{M \subseteq [n]} \pi(M) \lambda_i(A(\hat{b}^M), \hat{b}^M, b) . \quad (3)$$

The two formulas for payments in (2) and (3) must be equal:

$$\sum_{M \subseteq [n]} \pi(M) \lambda_i(A(\hat{b}^M), \hat{b}^M, b) = \sum_{M | i \notin M} \pi(M \cup \{i\}) \sum_{j \neq i} b_j(A(\hat{b}^M)) - \sum_{M | i \in M} \pi(M) \sum_{j \neq i} b_j(A(\hat{b}^M)) .$$

Since A may be any MIDR allocation function, the only way this can hold is when terms corresponding to each M are equal, i.e., for all i , M

$$\pi(M) \lambda_i(A(\hat{b}^M), \hat{b}^M, b) = \begin{cases} \pi(M \cup \{i\}) \sum_{j \neq i} b_j(A(\hat{b}^M)), & i \notin M \\ -\pi(M) \sum_{j \neq i} b_j(A(\hat{b}^M)) & i \in M . \end{cases} \quad (4)$$

To see that this is necessary, construct two allocation functions A and A' such that $b_j(A(\hat{b}^M)) = b_j(A'(\hat{b}^M))$ for all $M \neq \bar{M}$ and $b_j(A(\hat{b}^{\bar{M}})) = 0$. It immediately follows that if the reduction works for both A and A' , then (4) must hold for \bar{M} under A . Since \bar{M} is arbitrary, it follows that (4) must hold for all M .

The theorem immediately follows from the above equality. \blacksquare

Remark 1 Note that this theorem forbids some distributions $\pi(M)$ from being used to construct a single-call reduction — in particular, it requires that $\pi(M) > 0$ for all $M \subseteq [n]$, otherwise some payment $\lambda_i(\cdot)$ will be infinite for nontrivial allocation rules. For example, an obviously forbidden distribution is the one that never changes bids, i.e. the one with $\pi([n]) = 1$. This matches the intuition that a single-call mechanism must occasionally modify bids.

4.2 A Single-Call MIDR Reduction

We now give an explicit single-call reduction for MIDR allocation functions. Our reduction $\text{MIDRtoMech}(A, \gamma)$ (illustrated in Algorithm 2) is defined by the following resampling distribution $\bar{\pi}$ parameterized by a constant $\gamma \in (0, 1)$:

$$\bar{\pi}(M) = \gamma^{n-|M|} (1 - \gamma)^{|M|} \quad (5)$$

That is, each agent i is independently dropped from M with probability γ . Thus sampling from the distribution $\bar{\pi}$ is computationally easy. Following Theorem 4.1, we charge payments $\lambda_i(A(\hat{b}^M), \hat{b}^M, b) = c_i^{\bar{\pi}}(M) \sum_{j \neq i} b_j(A(\hat{b}^M))$ where

$$c_i^{\bar{\pi}}(M) = \begin{cases} -1, & i \in M \\ \frac{1-\gamma}{\gamma}, & i \notin M \end{cases}$$

Corollary 4.2 (of Theorem 4.1) *The mechanism*

$$\mathcal{M} = (\mathcal{A}, \{\mathcal{P}_i\}) = \text{MIDRtoMech}(A, \gamma)$$

calls A once and it satisfies truthfulness, individual rationality, and no positive transfers in an ex-post sense for all MIDR A .

4.3 Optimal Single-Call MIDR Reductions

We now prove that the construction $\text{MIDRtoMech}(A, \gamma)$ is optimal for the definitions of optimality given in Section 3. Theorem 4.1 implies that the bid-normalized payments will be

$$\sum_j \frac{\lambda_{ij}(b_j(A(\hat{b})), \hat{b}, b)}{b_j(A(\hat{b}))} = (n-1) c_i^{\pi}(M)$$

Thus, it is sufficient to optimize the variance as $\max_i \text{Var}_{M \sim \pi} c_i^{\pi}(M)$ and the worst-case as $\max_{i,M} |c_i^{\pi}(M)|$.

ALGORITHM 2: MIDRtoMech(A, γ) — A single-call reduction for MIDR allocation functions

input : MIDR allocation function A .

output: Truthful-in-expectation mechanism $\mathcal{M} = (\mathcal{A}, \{\mathcal{P}_i\})$.

1 Solicit bids b from agents;

2 **for** $i \in [n]$ **do**

with probability $1 - \gamma$

 Add agent i to set M ;

otherwise

 Drop agent i from M ;

3 Realize the outcome $A(\hat{b}^M)$;

4 Charge payments

$$\lambda_i(A(\hat{b}^M), \hat{b}^M, b) = \left(\sum_{j \neq i} b_j(A(\hat{b}^M)) \right) \times \begin{cases} -1, & i \in M \\ \frac{1-\gamma}{\gamma}, & i \notin M \end{cases};$$

4.3.1 Optimizing Risk vs. Precision

Theorem 4.3 *The reduction MIDRtoMech(A, γ) uniquely minimizes both the payment variance and the worst-case payment among all reductions that achieve a precision of at least $\alpha_P = (1 - \gamma)^n$.*

That is, for any other distribution π with precision $\pi([n]) \geq (1 - \gamma)^n$, the payment variance is larger, i.e.

$$\max_i \text{Var}_{M \sim \pi} c_i^\pi(M) > \max_i \text{Var}_{M \sim \bar{\pi}} c_i^{\bar{\pi}}(M) ,$$

and the worst-case payment is larger, i.e.

$$\max_{i, M} |c_i^\pi(M)| > \max_{i, M} |c_i^{\bar{\pi}}(M)| .$$

Proof: First we prove optimality for the worst-case payment $\max_{i, M} |c_i^\pi(M)|$ by contradiction. Assume that some distribution $\pi(M)$ does as well as $\bar{\pi}(M)$. Then it must be that $\max_{i, M} c_i^\pi(M) \leq \max_{i, M} c_i^{\bar{\pi}}(M)$ (the largest coefficient is not bigger), and $\pi([n]) \geq \bar{\pi}([n]) = \alpha_P$ (it respects the lower bound on precision). Since $\max c_i^{\bar{\pi}}(M) = \frac{1-\gamma}{\gamma}$, it must be that for all M and $i \notin M$,

$$\frac{\pi(M \cup \{i\})}{\pi(M)} \leq \max_{i, M} c_i^{\bar{\pi}}(M) = \frac{1-\gamma}{\gamma} = \frac{\bar{\pi}(M \cup \{i\})}{\bar{\pi}(M)} .$$

Therefore, for any bidder i , it must be that

$$\frac{\pi([n])}{\pi([n] \setminus \{i\})} \leq \frac{\bar{\pi}([n])}{\bar{\pi}([n] \setminus \{i\})} .$$

Since $\pi([n]) \geq \bar{\pi}([n])$, it follows that $\pi([n] \setminus \{i\}) \geq \bar{\pi}([n] \setminus \{i\})$. Repeating this argument, it follows by induction that $\pi(M) \geq \bar{\pi}(M)$ for any set M .

However, we also know that both $\pi(M)$ and $\bar{\pi}(M)$ are distributions so both have to sum to one over all M . Given that $\pi(M) \geq \bar{\pi}(M)$ for all M , this implies $\pi(M) = \bar{\pi}(M)$. Thus, $\bar{\pi}(M)$ is uniquely optimal.

Second, we argue that $\bar{\pi}$ optimizes the payment variance. The variance of bidder i 's payments is

$$\begin{aligned}\text{Var}_{M \sim \pi} c_i^\pi(M) &= \sum_{M \subseteq [n]} \pi(M) (c_i^\pi(M))^2 - \left(\sum_{M \subseteq [n]} \pi(M) c_i^\pi(M) \right)^2 \\ &= \sum_{M \subseteq [n]} \pi(M) (c_i^\pi(M))^2 - 0 \\ &= \sum_{M \subseteq [n] \setminus \{i\}} (\pi(M) + \pi(M \cup \{i\})) \frac{\pi(M \cup \{i\})}{\pi(M)}\end{aligned}$$

This is minimized when $\Pr(i \in M)$ is independent of other bidders (Lemma D.10), i.e. $\frac{\pi(M \cup \{i\})}{\pi(M)} = \frac{1-\gamma_i}{\gamma_i}$ for some constant γ_i . For such a distribution, the precision will be

$$\pi([n]) = \prod_i (1 - \gamma_i) .$$

It follows that the maximum variance is $\max_i \frac{1-\gamma_i}{\gamma_i}$, and it will only be minimized when $\gamma_i = \gamma_j$ for all $i \neq j$, which corresponds precisely to the distribution $\bar{\pi}$. ■

4.3.2 Optimizing Risk vs. Welfare

A natural optimization metric is the social welfare of \mathcal{A} (indeed, this was an open question from [BKS10] in the single-parameter setting).

Unfortunately, since MIDR allocation rules may generate negative utilities and remain MIDR under additive shifts of the valuation function, one can make the welfare approximation arbitrarily bad (indeed, even undefined) by subtracting a constant from each player's valuation. Thus, if valuation functions may be negative, we cannot meaningfully optimize the loss in social welfare.

However, when valuation functions are known to be nonnegative, then the following lemma shows that the worst-case welfare approximation is bounded:

Lemma 4.4 *The reduction $\text{MIDRtoMech}(A, \gamma)$ obtains an $\alpha_W = \min_i \Pr_\pi(i \in M) = 1 - \gamma$ approximation to the social welfare, and there is an allocation function A and bid b such that this bound is tight.*

The idea for the lower bound is that the sum of welfare of bidders in M cannot be lower at $A(\hat{b}^M)$ than at $A(\hat{b}^{[n]})$ because that would imply A did not maximize the social welfare of bidders in M at \hat{b}^M . The worst case scenario occurs when one player receives all the welfare. The proof is given in Appendix B.

Using this lemma, we can show that $\text{MIDRtoMech}(A, \gamma)$ is optimal:

Theorem 4.5 *The reduction $\text{MIDRtoMech}(A, \gamma)$ minimizes payment variance and worst-case payments among all reductions that achieve a welfare approximation of at least $\alpha_W = 1 - \gamma$.*

The proof is given in Appendix B.

4.3.3 Optimizing Risk vs. Revenue

The following lemma implies that a lower bound on the factor of approximation to revenue is equivalent to a lower bound on precision.

Lemma 4.6 *The reduction $\text{MIDRtoMech}(A, \gamma)$ obtains an $\alpha_\pi = \pi([n]) = (1 - \gamma)^n$ approximation to the revenue, and this is tight.*

Since Theorem 4.3 says that $\text{MIDRtoMech}(A, \gamma)$ optimizes payments with respect to precision, it similarly follows that it optimizes payments with respect to revenue:

Theorem 4.7 *The reduction $\text{MIDRtoMech}(A, \gamma)$ minimizes payment variance and the worst-case payment among all reductions that guarantee an $\alpha_R = (1 - \gamma)^n$ approximation to revenue.*

5 A Single-call application — PPC AdAuctions

Pay-per-click (PPC) AdAuctions are a prime example of mechanisms in which uncertainty can destroy truthfulness. There is a deep literature on truthful ad auctions, much of which makes a powerful assumption: the likelihood that a user clicks in any given setting is a commonly-held belief. In reality, this simply is not true. Auctioneers make their best effort to estimate the likelihood of a click; however, anecdotal evidence [Jab10] suggests that advertisers manipulate their bids according to the perceived accuracy of the auctioneer’s estimates. As we will illustrate in this section, even if the auctioneer’s estimates are good enough to (say) maximize welfare given the current bids, they are not sufficient to compute truthful prices. We show that single-call mechanisms can recover truthfulness in PPC ad auctions in spite of these conflicting beliefs.

In a standard PPC ad auction, n advertisers compete for $m \ll n$ slots. The value to an advertiser depends on the likelihood of a click, called the click-through-rate (CTR) c , and the value to the advertiser once the user has clicked, the value-per-click v . The expected value to an advertiser is thus cv . The auctioneer’s job is to assign advertisers to slots and compute per-click payments — bidders are only charged when a click occurs. Both tasks require knowing the CTRs for common objectives like welfare or revenue maximization, so the auctioneer must also maintain estimates of the CTRs, which we denote by c' .

Researchers generally acknowledge that, in reality, both c and v may depend arbitrarily on the outcome — they certainly depend on the quality and relevance of the particular ad being shown, but they also depend on where the ad is shown and on which other ads are shown nearby. However, for analytical tractability, the parameters c and v are often assumed to have a very restricted structure. We discuss two different structures to illustrate the pervasiveness of the problem caused by estimation error and to show how different single-call reductions may be applied.

Outcome-Independent Values and Separable CTRs In the ad auction literature, it is common to assume that a bidder’s value-per-click v_i is independent of the assignment and that the CTR is separable, that is, it takes the form $c = \alpha_j \beta_i$, where β_i depends only on the ad and α_j depends only on the slot $j \in [m]$ where the ad is shown. Unfortunately, even in this restricted setting, estimation errors may break the truthfulness of VCG prices. We give an example in Appendix A showing that even if the auctioneer’s estimates correctly identify the welfare-maximizing allocation, they may not yield truthful prices, even in the special case where $\beta_i = 1$.

In the language of allocations and payments, truthfulness is broken because the auctioneer only knows an estimate of A and thus does not have enough information to compute true VCG prices. However, once ads are shown, clicks may be measured, giving an unbiased estimate of bidders’ values. Unfortunately, this can only be done once — since the auctioneer only has one opportunity to show ads to the user, these unbiased estimates can only be measured under a single advertiser-slot assignment. Fortunately, *these unbiased estimates are exactly the information required to compute truthful payments using a single-call mechanism.*

Since a player’s bid b_i is merely its value-per-click v_i , this version of a PPC ad auction is a single-parameter domain and we can apply the result of [BKS10]. Their result says that we can turn any monotone allocation rule into a truthful-in-expectation mechanism — maximizing welfare subject to estimates α'_j and β'_i is a monotone allocation rule as long as the estimates α'_j have the same order as α_j (i.e. $\alpha'_{j_1} \geq \alpha'_{j_2}$ if $\alpha_{j_1} \geq \alpha_{j_2}$), so [BKS10] gives a truthful mechanism for almost any estimates:

Theorem 5.1 *Consider a single-parameter PPC auction with separable CTRs and let A^{PPC} be the allocation rule that maximizes welfare using estimated CTR parameters α'_j and β'_i , where the estimates α'_j are properly ordered. Then $\text{SPTtoMechBKS}(A^{PPC}, \gamma)$, the single-call reduction of [BKS10], gives a mechanism that is truthful in expectation and has expected welfare within a factor of $(1 - \gamma)^n$ of A^{PPC} .*

Outcome-Dependent Values and CTRs While most research uses single-parameter models for analytical tractability, an advertiser’s value-per-click v really depends on the advertiser-slot assignment chosen by the auctioneer as noted earlier. As in the preceding single-parameter setting, estimated CTRs are insufficient to guarantee truthfulness; however, the reduction of [BKS10] no-longer applies in such a multi-parameter domain — we show how our MIDR single-call reduction can be used to recover truthfulness.

To capture the dependence on the advertiser-slot assignment, we assume that a bidder’s CTR $c_{i,j}$ and value-per-click $v_{i,j}$ depend arbitrarily on both the bidder i and the slot j . Since the only allocation rules that have truthful prices in general multi-parameter domains are MIDR, we assume that the auctioneer can generate a MIDR allocation, specifically we assume the auctioneer can query an oracle to determine the allocation that maximizes the welfare of any set of bidders under the actual bid b (but not necessarily for an arbitrary bid b) and apply our MIDR reduction:

Theorem 5.2 *Consider a multi-parameter PPC auction where a bidder’s value-per-click $v_{i,j}$ depends on the bidder and the slot. Let A^{PPC} be an allocation rule that chooses the advertiser-slot assignment returned by the welfare-maximizing oracle described above. Then the mechanism $\text{MIDRtoMech}(A^{PPC}, \gamma)$ is truthful in expectation and approximates the welfare of A^{PPC} to within a factor of $(1 - \gamma)$.*

6 Single-parameter reductions

In this section, we characterize truthful reductions for single-parameter domains and show that the construction of [BKS10] is optimal. Theorem 6.1 characterizes all reductions that are truthful for an arbitrary monotone, bounded, single-parameter allocation function A . Our characterization is more general than the self resampling procedures described by Babaioff et al. and shows that a wide variety of probability measures may be used to construct a truthful reduction. Theorem 6.3 shows that the construction given in Babaioff et al. is optimal among such reductions for a fixed bound on the precision, welfare approximation, or revenue approximation of the reduction.

As in the MIDR setting, truthful payments give intuition for the structure of a single-call reduction. As noted in Section 2, payments are truthful if and only if they are given by the Archer-Tardos characterization:

$$p_i(b) = b_i A_i(b) - \int_0^{b_i} A_i(u, b_{-i}) du . \quad (6)$$

Loosely speaking, this says “charge i the value she receives minus what she would expect if she lowered her bid.” Thus, a single call reduction should, with some probability, lower agents’ bids to compute the value of allocation function at (u, b_{-i}) for $u \leq b_i$.

6.1 Characterizing Single-Call Reductions

For the sake of intuition, we start with the special case that the resampling measure μ_b has a nicely behaved density representation $f_b(\hat{b})$ (the resampling density) that is continuous in \hat{b} and b . The proof for arbitrary measures μ_b requires significant measure theory and is deferred until Appendix C.

Define the coefficients $c_i^f(\hat{b}, b)$ as $c_i^f(\hat{b}, b) = 1 - \frac{1}{b_i} \int_0^{b_i} \frac{f_{u, b-i}(\hat{b})}{f_b(\hat{b})} du$ when $b_i \neq 0$, and to be 0 when $b_i = 0$. We characterize truthful reductions as follows:

Theorem 6.1 *A normalized single-parameter reduction $(f, \{\lambda_i\})$ for the set of all monotone bounded single-parameter allocation functions satisfies truthfulness, individual rationality and no positive transfers in an ex-post sense if and only if the following conditions are met:*

1. *The resampling density f_b is such that the single-call mechanism's randomized allocation procedure $\mathcal{A}_i(b)$ is monotone in expectation, i.e., for all agents i , for all b , and $b'_i \geq b_i$, $\mathbf{E}_{\hat{b} \sim f_b}[\mathcal{A}_i(b'_i, b_{-i})] \geq \mathbf{E}_{\hat{b} \sim f_b}[\mathcal{A}_i(b)]$. (See below.)*
2. *The resampling density f_b is such that $f_b(\hat{b}) \neq 0$ if $\int_0^{b_i} f_{u, b-i}(\hat{b}) du \neq 0$.⁹*
3. *The payment functions $\lambda_i(A(\hat{b}), \hat{b}, b)$ satisfy: $\lambda_i(A(\hat{b}), \hat{b}, b) = b_i c_i^f(\hat{b}, b) \mathcal{A}_i(\hat{b})$ almost surely, i.e. for all \hat{b} except possibly a set with probability zero under f_b .*

Proof: (See Appendix C for the proof when μ_b is an arbitrary measure.)

Necessity. The first condition, that \mathcal{A} must be monotone in expectation, follows directly from Archer-Tardos characterization of truthful allocation functions. The second and third conditions, as we prove below, are necessary for the expected payment to take the form required by the Archer-Tardos characterization.

The allocation function \mathcal{A} is a single-parameter allocation function, so the Archer-Tardos characterization gives truthful prices if they exist:

$$\begin{aligned} \mathbf{E}[\mathcal{P}_i] &= b_i \mathbf{E}_{\hat{b} \sim f_b}[\mathcal{A}_i(b)] - \int_0^{b_i} \mathbf{E}_{\hat{b} \sim f_{u, b-i}}[\mathcal{A}_i(u, b_{-i})] du \\ &= b_i \mathbf{E}_{\hat{b} \sim f_b}[\mathcal{A}_i(\hat{b})] - \int_0^{b_i} \mathbf{E}_{\hat{b} \sim f_{u, b-i}}[\mathcal{A}_i(\hat{b})] du \\ &= b_i \int_{\hat{b} \in \mathbb{R}^n} \mathcal{A}_i(\hat{b}) f_b(\hat{b}) d\hat{b} - \int_0^{b_i} \int_{\hat{b} \in \mathbb{R}^n} \mathcal{A}_i(\hat{b}) f_{u, b-i}(\hat{b}) d\hat{b} du . \end{aligned}$$

Rearranging, where changing the order of integration may be justified by Tonelli's theorem, gives

$$\mathbf{E}[\mathcal{P}_i] = \int_{\hat{b} \in \mathbb{R}^n} f_b(\hat{b}) b_i \mathcal{A}_i(\hat{b}) \left(1 - \frac{1}{b_i} \int_0^{b_i} \frac{f_{u, b-i}(\hat{b})}{f_b(\hat{b})} du \right) d\hat{b} .$$

By construction, we can express the expected price as

$$\mathbf{E}[\mathcal{P}_i] = \int_{\hat{b} \in \mathbb{R}^n} f_b(\hat{b}) \lambda_i(A(\hat{b}), \hat{b}, b) d\hat{b} .$$

Thus truthfulness in expectation necessarily implies

$$\int_{\hat{b} \in \mathbb{R}^n} f_b(\hat{b}) \lambda_i(A(\hat{b}), \hat{b}, b) d\hat{b} = \int_{\hat{b} \in \mathbb{R}^n} f_b(\hat{b}) b_i \mathcal{A}_i(\hat{b}) \left(1 - \frac{1}{b_i} \int_0^{b_i} \frac{f_{u, b-i}(\hat{b})}{f_b(\hat{b})} du \right) d\hat{b} . \quad (7)$$

⁹This condition effectively requires $c_i^f(\hat{b}, b)$ to be finite.

Note that proving the necessity of condition three in the theorem is equivalent to proving that the integrands in the LHS and the RHS of (7) are equal almost everywhere. That is, we have to show that the only way for Equation (7) to hold for all monotone bounded A is when the integrands are equal almost everywhere. To show this, it is sufficient to show that Equation (7) must still hold if we restrict the range of integration to an arbitrary rectangular parallelepiped (hence forth called as rectangle) $\mathbb{S} \subseteq \mathbb{R}^n$ (see why this is enough in Appendix C for a more general setting), that is, it is sufficient to show that for all rectangles $\mathbb{S} \subseteq \mathbb{R}^n$

$$\int_{\hat{b} \in \mathbb{S}} f_b(\hat{b}) \lambda_i(A(\hat{b}), \hat{b}, b) d\hat{b} = \int_{\hat{b} \in \mathbb{S}} f_b(\hat{b}) b_i A_i(\hat{b}) \left(1 - \frac{1}{b_i} \int_0^{b_i} \frac{f_{u, b-i}(\hat{b})}{f_b(\hat{b})} du \right) d\hat{b} . \quad (8)$$

Showing (8) would be straight-forward if we are given that (7) holds for *all* A — we could take any A and make it zero for all points not in \mathbb{S} , and then (7) immediately implies (8). However (7) is guaranteed to be true only for monotone bounded A , since those are the allocation functions that could possibly be input to our reduction. To see that it is still true when (7) is only guaranteed for monotone bounded A , define the function $1_{\mathbb{S}}(\hat{b})$ as

$$1_{\mathbb{S}}(\hat{b}) = \begin{cases} 1, & \hat{b} \in \mathbb{S} \\ 0, & \text{otherwise.} \end{cases}$$

Observe that $1_{\mathbb{S}}$ can be written as $1_{\mathbb{S}}(\hat{b}) = 1_{\mathbb{S}}^+(\hat{b}) - 1_{\mathbb{S}}^-(\hat{b})$ where $1_{\mathbb{S}}^+$ and $1_{\mathbb{S}}^-$ are both $\{0, 1\}$, monotone functions. Moreover, the functions $A^+(b) = 1_{\mathbb{S}}^+(b)A(b)$ and $A^-(b) = 1_{\mathbb{S}}^-(b)A(b)$ are also monotone, and they agree with A on \mathbb{S} . If we plug A^+ and A^- into (7) and subtract the results, we get precisely (8). Thus condition three is necessary.

For the necessity of condition two, note that if it were not to hold, the coefficients c_i^f will become $-\infty$, and hence the payments as defined in condition three will not be finite. Clearly finiteness of payments is a requirement.

This proves that all three conditions in the theorem are necessary for truthfulness.

Sufficiency. We now show that the three stated conditions are sufficient. In a single-parameter setting, for a mechanism to be truthful, we need the allocation function to be monotone in expectation and the payment function to satisfy the Archer-Tardos payment functions. Condition one guarantees that the allocation function output by the single-call reduction is a monotone in expectation allocation function. It remains to show that the second and third conditions result in payments that agree with Archer-Tardos payments. Given condition two, finiteness of payments as defined in condition three is satisfied. All we need to show is that under the formula of $\lambda_i(A(\hat{b}), \hat{b}, b)$ described in condition three, the single-call payments match in expectation with Archer-Tardos payments, i.e., (7) holds. Since $c_i^f(\hat{b}, b) = 1 - \frac{1}{b_i} \int_0^{b_i} \frac{f_{u, b-i}(\hat{b})}{f_b(\hat{b})} du$, taking

$$\lambda_i(A(\hat{b}), \hat{b}, b) = b_i c_i^f(\hat{b}, b) A_i(\hat{b}) \quad a.s.$$

trivially satisfies (7), implying that the reduction is truthful. ■

Unfortunately, our assumption that μ_b has a density representation is unreasonable. Most significantly, one would expect $\hat{b} = b$ with some nonzero probability, implying that μ_b would have at least one atom for most interesting distributions. In particular, the distribution used in the BKS transformation has such an atom, so it cannot be analyzed in this fashion.

To handle general measures μ_b we apply the same ideas using tools from measure theory. A full proof is given in Appendix C.

ALGORITHM 3: $\text{SPtoMechBKS}(A, \gamma)$ — The BKS reduction for single-parameter domains

input : Bounded, monotone allocation function A .

output: Truthful-in-expectation mechanism $\mathcal{M} = (\mathcal{A}, \{\mathcal{P}_i\})$.

1 Solicit bids b from agents;

2 **for** $i \in [n]$ **do**

with probability $1 - \gamma$

 Set $\hat{b}_i = b_i$;

otherwise

 Sample x_i uniformly at random from $[0, \hat{b}_i]$;

 Set $\hat{b}_i = b_i x_i^{\frac{1}{1-\gamma}}$;

3 Realize the outcome $A(\hat{b})$;

4 Charge payments

$$\lambda_i(A(\hat{b}^M), \hat{b}^M, b) = b_i A_i(\hat{b}) \times \begin{cases} 1, & \hat{b}_i = b_i; \\ \frac{1-\gamma}{\gamma}, & \hat{b}_i < b_i; \end{cases}$$

6.2 The BKS Reduction for Positive Types

The central construction of Babaioff, Kleinberg, and Slivkins [BKS10] is a reduction for scenarios where bidders have positive types.¹⁰

Their resampling procedure (implicitly defining μ_b) is described Algorithm 3. In the language of our characterization, the coefficients c_i^{BKS} are

$$c_i^{BKS}(\hat{b}, b) = \begin{cases} 1, & \hat{b}_i = b_i \\ 1 - \frac{1}{\gamma} & \text{otherwise.} \end{cases}$$

They proved that $\text{SPtoMechBKS}(A, \gamma)$ is truthful. This fact can be easily derived from Theorem 6.1:

Theorem 6.2 (Babaioff, Kleinberg, and Slivkins 2010.) *For all monotone, bounded, single-parameter allocation rules A , the single-call mechanism given by $\text{SPtoMechBKS}(A, \gamma)$ satisfies truthfulness and no positive transfers in an ex-post sense and is ex-post universally individually rational.*

6.3 Optimal Single-Call Reductions

Analogous to our MIDR construction, we show that, the BKS construction for positive types is optimal with respect to precision, welfare, and revenue as defined in Section 3 (other type spaces are discussed in Appendix C). Using our characterization from Theorem 6.1, the bid-normalized payments we wish to optimize will be

$$\sum_j \frac{\lambda_{ij}(b_j(A(\hat{b})), \hat{b}, b)}{b_j(A(\hat{b}))} = \frac{c_i^\mu(\hat{b}, b) b_i A_i(\hat{b})}{b_i A_i(\hat{b})} = c_i^\mu(\hat{b}, b) .$$

Thus, optimizing variance of normalized payments is equivalent to optimizing $\max_i \text{Var}_{\hat{b} \sim \mu_b} c_i^\mu(\hat{b}, b)$, and optimizing the worst-case normalized payment is equivalent to optimizing $\sup_{i, \hat{b}} |c_i^\mu(\hat{b}, b)|$.

¹⁰They also give a reduction that applies to more general type spaces, but we do not state it here.

For this section, we make a “nice distribution” assumption that for any $u \neq b_i$, $\Pr(\hat{b}_i = u|b) = 0$. That is, if we compute the marginal distribution of \hat{b}_i , the only bid \hat{b}_i that has an atom is b_i (other bids only have positive density). We handle the general case in the full proofs in Appendix D.

Our main result is that the BKS transformation is optimal:

Theorem 6.3 *The single-call reduction $\text{SPtoMechBKS}(A, \gamma)$ optimizes the variance of bid-normalized payments and the worst-case bid-normalized payment*

for every b subject to a lower bound $\alpha = (1 - \gamma)^n \in (\frac{1}{e}, 1)$ on the precision, the welfare approximation, or the revenue approximation.

To prove Theorem 6.3, we first show that the three metrics we study are equivalent for interesting reductions in the single parameter setting:

Lemma 6.4 *For $\alpha > \frac{1}{e}$ and $n \geq 2$, a reduction that optimizes the variance of normalized payments or the maximum normalized payment subject to a precision constraint of $\Pr(\hat{b} = b|b) \geq \alpha$ also optimizes the maximum payment subject to a welfare or revenue approximation of α .*

Proof: (Sketch. The full proof is in Appendix D.) Consider the following allocation function:

$$A_i(b) = \begin{cases} 1, & b \geq \bar{b} \\ 0, & \text{otherwise.} \end{cases}$$

Intuitively, a reduction should not resample to higher bids because Archer-Tardos payments do not depend on higher bids, and hence no useful information is obtained through raising bids. However, if a reduction never raises bids (i.e. $\Pr(\hat{b} \leq b|b) = 1$), then the welfare and revenue of a single-call reduction will both be precisely $\Pr(\hat{b} = b|b)$ if we consider the above mentioned A at a bid of \bar{b} . ■

Thus, to prove Theorem 6.3, it is sufficient to prove that the BKS reduction optimizes precision.

Theorem 6.5 *The single-call reduction $\text{SPtoMechBKS}(A, \gamma)$ optimizes the variance of normalized payments and the worst-case normalized payment among reductions with a precision of at least $\alpha_P = (1 - \gamma)^n > \frac{1}{e}$.*

Proof: (Sketch. The full proof is in Appendix D.) When $\Pr(\hat{b} = b|b)$ is large, the mechanism extracts a modest payment from i when $\hat{b}_i = b_i$ and pays a large rebate otherwise. Thus, we bound $\inf_{\hat{b}_i} c_i^\mu(\hat{b}, b)$. Let $\pi^\mu(M, b)$ be the probability (given b) that $\hat{b}_i = b_i$ for all $i \in M$ and $\hat{b}_i < b_i$ for all $i \notin M$. Then the key step is to prove the following lower bound on $\inf_{\hat{b}} c_i^\mu$:

$$\inf_{\hat{b}} c_i^\mu(\hat{b}, b) \leq -\frac{\pi^\mu(M \cup \{i\}, b)}{\pi^\mu(M, b)}.$$

Notably, this bound takes the same form as the truthful payment coefficients for MIDR reductions. Applying the same logic as Theorem 4.3 shows that the BKS transformation is optimal. ■

7 Acknowledgments

First, we would like to thank Kamal Jain for enlightening us about the problems that can arise in pay-per-click advertising auctions and for suggesting [BKS10] as a possible solution.

Second, we would like to thank many people who have provided invaluable feedback and suggestions. In particular, we would like to thank Alex Slivkins for his substantial help in the revising process, Robert Kleinberg for his technical suggestions, and Christos Papadimitriou for his helpful comments.

References

- [AT01] A. Archer and É. Tardos. Truthful mechanisms for one-parameter agents. In *FOCS '01: Proceedings of the 42nd IEEE symposium on Foundations of Computer Science*, page 482, Washington, DC, USA, 2001. IEEE Computer Society.
- [BBNS08] Moshe Babaioff, Liad Blumrosen, Moni Naor, and Michael Schapira. Informational overhead of incentive compatibility. In *EC '08: Proceedings of the 9th ACM conference on Electronic commerce*, pages 88–97, New York, NY, USA, 2008. ACM.
- [BKS10] Moshe Babaioff, Robert D. Kleinberg, and Aleksandrs Slivkins. Truthful mechanisms with implicit payment computation. In *EC '10: Proceedings of the 11th ACM conference on Electronic commerce*, pages 43–52, New York, NY, USA, 2010. ACM.
- [BSS09] Moshe Babaioff, Yogeshwer Sharma, and Aleksandrs Slivkins. Characterizing truthful multi-armed bandit mechanisms: extended abstract. In *EC '09: Proceedings of the tenth ACM conference on Electronic commerce*, pages 79–88, New York, NY, USA, 2009. ACM.
- [DD09] Shahar Dobzinski and Shaddin Dughmi. On the power of randomization in algorithmic mechanism design. In *FOCS*, pages 505–514, 2009.
- [DK09] Nikhil R. Devanur and Sham M. Kakade. The price of truthfulness for pay-per-click auctions. In *EC '09: Proceedings of the tenth ACM conference on Electronic commerce*, pages 99–106, New York, NY, USA, 2009. ACM.
- [DR10] Shaddin Dughmi and Tim Roughgarden. Black-box randomized reductions in algorithmic mechanism design. In *Proceedings of the 51st IEEE symposium on Foundations of Computer Science*, pages 775–784, 2010.
- [Har11] Jason D. Hartline. Approximation in economic design, 2011. Draft.
- [HS01] J. Hershberger and S. Suri. Vickrey prices and shortest paths: What is an edge worth? In *Proceedings of the 42nd IEEE symposium on Foundations of Computer Science*, FOCS '01, Washington, DC, USA, 2001. IEEE Computer Society.
- [HS02] John Hershberger and Subhash Suri. Erratum to “Vickrey pricing and shortest paths: What is an edge worth?”. In *Proceedings of the 43rd Symposium on Foundations of Computer Science*, FOCS '02, Washington, DC, USA, 2002. IEEE Computer Society.
- [HSB07] John Hershberger, Subhash Suri, and Amit Bhosle. On the difficulty of some shortest path problems. *ACM Trans. Algorithms*, 3:5:1–5:15, February 2007.
- [Jab10] David Jabas. Private communication, 2010.
- [Lah10] Sébastien Lahaie. Stability and incentive compatibility in a kernel-based combinatorial auction. In *AAAI*, 2010.
- [Mye81] Roger B. Myerson. Optimal Auction Design. *Mathematics of Operations Research*, 6(1):58–73, 1981.
- [NR01] Noam Nisan and Amir Ronen. Algorithmic mechanism design. *Games and Economic Behavior*, 35(1-2):166–196, April 2001.

- [NR07] Noam Nisan and Amir Ronen. Computationally feasible VCG mechanisms. *J. Artif. Intell. Res. (JAIR)*, 29:19–47, 2007.
- [Rob79] Kevin Roberts. The characterization of implementable social choice rules. In *Aggregation and Revelation of Preferences*, J-J.Laffont (ed.), North Holland Publishing Company., 1979.

A A PPC Auction Example

The following example illustrates how the welfare optimal assignment may be robust to inaccuracies in the CTR estimates c' but the truthful payments are quite fragile.

Example 1 Consider a 2-slot, 2-advertiser setting with CTRs c_j and bids b_i . Assume that $b_1 > b_2$ and $c_1 > c_2$, so that the welfare-optimizing assignment is to assign ad-1 to slot-1 and ad-2 to slot-2, i.e.,

$$c_1 b_1 + c_2 b_2 \geq c_1 b_2 + c_2 b_1 . \quad (9)$$

The auctioneer wishes to optimize welfare, so he uses c'_j to implement the VCG allocation. It is quite plausible that maximizing welfare w.r.t c'_j results in the same welfare maximizing allocation, namely given (9), it is not unreasonable to assume that the following is true if the auctioneer's estimates are good enough:

$$c'_1 b_1 + c'_2 b_2 \geq c'_1 b_2 + c'_2 b_1 .$$

However, we will show that this is not enough to guarantee truthfulness.

We show that advertiser-1 may have an incentive to lie. According to the estimates c'_j , The expected VCG payment should be $c'_1 b_2 - c'_2 b_2$. Since advertiser 1 will only be charged when he actually receives a click, the price-per-click charged will be

$$p_1 = \frac{1}{c'_1} [c'_1 b_2 - c'_2 b_2] .$$

and the expected utility to bidder i will be

$$u_1 = c_1 \left(b_1 - \frac{1}{c'_1} [c'_1 b_2 - c'_2 b_2] \right) ,$$

where the extra c_1 gets multiplied because the utility is non-zero only upon a click, which happens with probability c_1 .

Now, for example, let the inaccurate c'_j be as follows: $c'_1 = \alpha c_1$, $c'_2 = c_2$ where $\alpha > 1$. Notice that in this example we always have

$$\alpha c_1 b_1 + c_2 b_2 \geq \alpha c_1 b_2 + c_2 b_1$$

and thus the mechanism will always maximize welfare in spite of the estimation errors.

The utility of advertiser-1 will be

$$u_1 = c_1 \left(b_1 - \frac{1}{\alpha c_1} [\alpha c_1 b_2 - c_2 b_2] \right) .$$

Now, suppose advertiser-1 decides to lie and bid zero, he gets the second slot, pays zero, and gets utility of $c_2 b_1$. Lying is clearly profitable if

$$c_2 b_1 > c_1 \left(b_1 - \frac{1}{\alpha c_1} [\alpha c_1 b_2 - c_2 b_2] \right) .$$

Rearranging, lying is profitable if

$$\alpha c_1 b_1 + c_2 b_2 < \alpha c_1 b_2 + \alpha c_2 b_1 \quad (10)$$

It is quite possible that lying might be profitable, that is inequality (10) holds true. For example, if $c_1 = 0.1$, $c_2 = 0.09$, $b_1 = 1.1$, $b_2 = 1$, and $\alpha = 1.1$, payments computed using c'_j are nontruthful, even though the mechanism always picks the welfare-maximizing assignment for any $\alpha > 1$.

B Optimality Proofs for MIDR Reductions

B.1 Optimizing Social Welfare

Lemma B.1 (Restatement of Lemma 4.4) *The reduction $\text{MIDR}_{\text{toMech}}(A, \gamma)$ obtains an $\alpha_\pi = \min_i \Pr_\pi(i \in M)$ approximation to the social welfare, and there is an allocation function A and bid b such that this bound is tight.*

Proof: The expected social welfare of the single-call mechanism, where the expectation is over the randomness in the resampling function is given by $\mathbf{E} \left[\sum_{j \in [n]} b_j(\mathcal{A}(b)) \right]$. We now prove the required lower bound on this quantity.

$$\begin{aligned} \mathbf{E} \left[\sum_{j \in [n]} b_j(\mathcal{A}(b)) \right] &= \sum_{j \in [n]} \sum_{M \subseteq [n]} \pi(M) b_j(A(\hat{b}^M)) = \sum_{M \subseteq [n]} \pi(M) \sum_{j \in [n]} b_j(A(\hat{b}^M)) \\ &\geq \sum_{M \subseteq [n]} \pi(M) \sum_{j \in M} b_j(A(\hat{b}^M)) \\ &\geq \sum_{M \subseteq [n]} \pi(M) \sum_{j \in M} b_j(A(\hat{b}^{[n]})) \\ &= \sum_{j \in [n]} \Pr_\pi(j \in M) b_j(A(\hat{b}^{[n]})) \\ &\geq \left(\min_{j \in [n]} \Pr_\pi(j \in M) \right) \sum_{j \in [n]} b_j(A(\hat{b}^{[n]})) \\ &= \left(1 - \max_{j \in [n]} \Pr_\pi(j \notin M) \right) \sum_{j \in [n]} b_j(A(\hat{b}^{[n]})) . \end{aligned}$$

Finally, we observe that this is tight. Consider a valuation and allocation function pair for which, every agent other than some agent j has a zero value for every outcome, and agent j has a non-zero value only for those outcomes that were chosen taking j into consideration, i.e., :

$$b_k(A(\hat{b}^M)) = \begin{cases} 0, & k \neq j \\ 0, & j \notin M \\ 1, & \text{otherwise} \end{cases}$$

When $j = \operatorname{argmax}_{k \in [n]} \Pr_\pi(k \notin M)$, the preceding bound is tight. ■

Lemma B.2 *Let π be a distribution such that $\max_{i,M} c_i^\pi(M) < \max_{i,M} c_i^{\bar{\pi}}(M)$. Then*

$$\max_i \Pr_\pi(i \notin M) > \max_i \Pr_{\bar{\pi}}(i \notin M) .$$

Proof:

Let $\bar{c} = \max_{i,M} c_i^{\bar{\pi}}(M)$. Note that for all $M|i \notin M$, $c_i^{\bar{\pi}}(M) = \frac{\bar{\pi}(M \cup \{i\})}{\bar{\pi}(M)} = \bar{c}$. It follows by algebra that $\frac{\sum_{M|i \notin M} \bar{\pi}(M \cup \{i\})}{\sum_{M|i \notin M} \bar{\pi}(M)} = \bar{c}$ and therefore by the conditions of the lemma

$$\max_{i,M} c_i^\pi(M) < \frac{\sum_{M|i \notin M} \bar{\pi}(M \cup \{i\})}{\sum_{M|i \notin M} \bar{\pi}(M)} . \quad (11)$$

Next we have,

$$\max_{i,M} c_i^\pi(M) \geq \max_M c_i^\pi(M) \geq \max_{M|i \notin M} \frac{\pi(M \cup \{i\})}{\pi(M)} \geq \frac{\sum_{M|i \notin M} \pi(M \cup \{i\})}{\sum_{M|i \notin M} \pi(M)} \quad (12)$$

Combining (11) and (12) gives

$$\frac{\sum_{M|i \notin M} \pi(M \cup \{i\})}{\sum_{M|i \notin M} \pi(M)} < \frac{\sum_{M|i \notin M} \bar{\pi}(M \cup \{i\})}{\sum_{M|i \notin M} \bar{\pi}(M)} \quad (13)$$

Note that since π and $\bar{\pi}$ are probability distributions, the sum of the numerator and denominator of both the LHS and the RHS of (13) equals 1. Thus, it immediately follows that the denominator of the LHS is larger than the denominator of the RHS, i.e.,

$$\sum_{M|i \notin M} \pi(M) > \sum_{M|i \notin M} \bar{\pi}(M) \quad (14)$$

Inequality (14) when restated, reads as

$$\Pr_\pi(i \notin M) > \Pr_{\bar{\pi}}(i \notin M) .$$

But since the above inequality is true for all i , and the RHS of the above inequality is the same for all i (namely the parameter μ by which the reduction is parametrized), the statement of the lemma follows. ■

Theorem B.3 *(Restatement of Theorem 4.5.) The reduction $\text{MIDR}_{\text{toMech}}(A, \gamma)$ minimizes payment variance and the worst-case payment among all reductions that achieve a welfare approximation of at least $\alpha_W = 1 - \gamma$.*

Proof: By Lemma B.1, the worst case loss in social welfare of a distribution π is given by

$$1 - \alpha_\pi = \max_i \Pr_\pi(i \notin M) .$$

For worst-case payments, the contrapositive of Lemma B.2 precisely says that if $1 - \alpha_\pi \leq 1 - \alpha_{\bar{\pi}}$, then the largest payment $\max_{M,i} c_i^\pi(M) \geq \max_{M,i} c_i^{\bar{\pi}}(M)$, thus proving that any other reduction will be worse.

For payment variance, arguing along the lines of Theorem 4.1 again says that variance will be minimized when π is an independent distribution and $\Pr(i \in M)$ is the same for all i . Since $\bar{\pi}$ is precisely the distribution that does this, it follows that it is optimal. ■

B.2 Optimizing Revenue

Lemma B.4 (Restatement of Lemma 4.6) *The reduction $\text{MIDRtoMech}(A, \gamma)$ obtains an $\alpha_\pi = \pi([n])$ approximation to the revenue, and this is tight.*

Proof: For any b with non-negative valuations, the revenue under a single call reduction will be

$$\begin{aligned} \sum_{i \in [n]} \mathbf{E}[P_i] &= \sum_{i \in [n]} \sum_{M \subseteq [n]} \pi(M) \left(\sum_{k \neq i} b_k(A(\hat{b}^{M \setminus \{i\}})) - \sum_{k \neq i} b_k(A(\hat{b}^M)) \right) \\ &\geq \pi([n]) \sum_{i \in [n]} \left(\sum_{k \neq i} b_k(A(\hat{b}^{[n] \setminus \{i\}})) - \sum_{k \neq i} b_k(A(\hat{b}^{[n]})) \right) \end{aligned}$$

where $\sum_{i \in [n]} \left(\sum_{k \neq i} b_k(A(\hat{b}^{[n] \setminus \{i\}})) - \sum_{k \neq i} b_k(A(\hat{b}^{[n]})) \right)$ is the revenue generated by A under VCG prices.

Thus, any distribution $\pi(M)$ gives an $\alpha = \pi([n])$ approximation to the revenue.

To see that this is tight, consider the following allocation function:

$$b_i(A(\hat{b}^M)) = \begin{cases} \frac{1}{n} & M = [n] \\ \frac{1}{n-1} & i \in M \text{ but } M \neq [n] \\ 0, & \text{otherwise.} \end{cases}$$

The revenue under VCG prices is $\sum_{i \in [n]} \left(\sum_{k \neq i} b_k(A(\hat{b}^{[n] \setminus \{i\}})) - \sum_{k \neq i} b_k(A(\hat{b}^{[n]})) \right)$, which is $n \left(\frac{n-1}{n-1} - \frac{n-1}{n} \right) = 1$.

Under any single-call reduction, the revenue will be given by

$$\begin{aligned} \sum_{i \in [n]} \mathbf{E}[P_i] &= \sum_{i \in [n]} \sum_{M \subseteq [n]} \pi(M) \left(\sum_{k \neq i} b_k(A(\hat{b}^{M \setminus \{i\}})) - \sum_{k \neq i} b_k(A(\hat{b}^M)) \right) \\ &= \sum_{i \in [n]} \pi([n]) \left(\sum_{k \neq i} b_k(A(\hat{b}^{[n] \setminus \{i\}})) - \sum_{k \neq i} b_k(A(\hat{b}^{[n]})) \right) \\ &= \sum_{i \in [n]} \pi([n]) \left(1 - \frac{n-1}{n} \right) \\ &= \pi([n]) . \end{aligned}$$

■

C Characterizing Reductions for Single-Parameter Domains

In this section we characterize truthful single-call reductions for single-parameter domains that use arbitrary measures μ_b . We refer the reader to Section E for some background and definitions from measure theory.

Before we begin, we must formalize some properties of the functions A and the measures μ_b . The following assumptions would typically be implicit in Algorithmic Mechanism Design; however, it is necessary that they be formalized for some of the tools in our proof. We assume the following:

1. Any allocation function A that the reduction receives as input (as a black box) is a Borel measurable function, i.e., each of the A_i 's as a function from $\mathbb{R}^n \rightarrow \mathbb{R}_+$ is a bounded Borel measurable function.
2. For every b , the resampling measure $\mu_b(\cdot)$ is a Borel probability measure.
3. The function mapping the bid b to the resampling measure $\mu_b(\cdot)$ is measurable w.r.t to the Borel σ -algebra on the space of Borel probability measures over \mathbb{R}^n .

First, we use the measure $\mu_b(\cdot)$ to define a signed measure $\nu_{b,i}(B) = b_i \mu_b(B) - \int_0^{b_i} \mu_{u,b-i}(B) du$ which has the property:

$$\int_{\hat{b} \in \mathbb{R}^n} A_i(\hat{b}) d\nu_{b,i} = b_i \mathbf{E}_{\hat{b} \sim \mu_b} [A_i(\hat{b})] - \int_0^{b_i} \mathbf{E}_{\hat{b} \sim \mu_{u,b-i}} [A_i(\hat{b})] du ,$$

that is, integrating A_i with respect to $\nu_{b,i}$ is equivalent to computing the Archer-Tardos prices.

Lemma C.1 *The function $\nu_{b,i}(B) = b_i \mu_b(B) - \int_0^{b_i} \mu_{u,b-i}(B) du$ is a finite signed measure satisfying*

$$\int_{\hat{b} \in \mathbb{R}^n} A_i(\hat{b}) d\nu_{b,i} = b_i \mathbf{E}_{\hat{b} \sim \mu_b} [A_i(\hat{b})] - \int_0^{b_i} \mathbf{E}_{\hat{b} \sim \mu_{u,b-i}} [A_i(\hat{b})] du$$

for any bounded A_i

Proof: First, we show that $\nu_{b,i}(B)$ is a finite signed measure. Since μ_b is a probability measure, we have $\mu_b(B) \leq 1$ for all B . Thus, $\nu_{b,i}(B)$ is well-defined and finite for all Borel sets B (note that the integral is well defined by our assumptions on the measurability of μ_b). From this it is easy to see that $\nu_{b,i}(\emptyset) = 0$ because $\mu_b(\emptyset) = 0$. It remains to show countable additivity, i.e. $\sum_{k=1}^{\infty} \nu_{b,i}(B_k) = \nu_{b,i}(\cup_k B_k)$, which follows because integrals obey countable additivity for nonnegative functions (see Fact E.7):

$$\begin{aligned} \sum_{k=1}^{\infty} \nu_{b,i}(B_k) &= \sum_{k=1}^{\infty} \left(b_i \mu_b(B_k) - \int_0^{b_i} \mu_{u,b-i}(B_k) du \right) = \sum_{k=1}^{\infty} b_i \mu_b(B_k) - \int_0^{b_i} \sum_{k=1}^{\infty} \mu_{u,b-i}(B_k) du \\ &= b_i \mu_b(\cup_k B_k) - \int_0^{b_i} \mu_{u,b-i}(\cup_k B_k) du = \nu_{b,i}(\cup_k B_k) . \end{aligned}$$

Second, we show from first-principles that integrating A_i with respect to $\nu_{b,i}$ is equivalent to calculating the Archer-Tardos prices for A_i . We begin by showing this equality for characteristic functions over Borel measurable sets. The proof for more general functions (in our case A_i) can be built-up from characteristic functions precisely as in the definition of an integral, so we omit it (see Definition 21). Let 1_B be the characteristic function of a Borel measurable set. By definition of an integral, $\int 1_X d\nu = \nu(X)$, and plugging in we observe the desired equality:

$$\begin{aligned} b_i \mathbf{E}_{\hat{b} \sim \mu_b} [1_B(\hat{b})] - \int_0^{b_i} \mathbf{E}_{\hat{b} \sim \mu_{u,b-i}} [1_B(\hat{b})] du &= b_i \mu_b(B) - \int_0^{b_i} \mu_{u,b-i}(B) du \\ &= \int_{\hat{b} \in \mathbb{R}^n} 1_B(\hat{b}) d\nu_{b,i} . \end{aligned}$$

■

The general version of the characterization theorem shows that the payment functions precisely correspond to the density function $\rho_{b,i}^{\mu}(\hat{b})$ relating $\nu_{b,i}$ to μ_b (i.e. the Radon-Nikodym derivative of $\nu_{b,i}$ with

respect to μ_b — its existence is guaranteed by the absolute continuity that figures in the characterization theorem C.2 below). In this setting, we can equivalently define the associated coefficients $c_i^\mu(\hat{b}, b)$ as the function that satisfies

$$b_i c_i^\mu(\hat{b}, b) = \rho_{b,i}^\mu(\hat{b}) .$$

Theorem C.2 (Characterizing single-call reductions) (Generalization of Theorem 6.1.)

A single-call single-parameter reduction $(\mu, \{\lambda_i\})$ for the set of all monotone bounded single-parameter allocation functions satisfies truthfulness, individual rationality, and no positive transfers in expectation if and only if the following conditions are met:

1. The distribution μ is such that for all monotone, locally bounded A , the randomized allocation procedure $\mathcal{A}_i(b)$ is monotone in expectation, i.e., for all agents i , for all b , and $b'_i \geq b_i$, $\mathbf{E}[\mathcal{A}_i(b)] \leq \mathbf{E}[\mathcal{A}_i(b', b_{-i})]$ (see Lemma C.4 for further discussion).
2. For all i , and for all Borel measurable sets B , the measure $\mu_b(B) \neq 0$ if $\int_0^{b_i} \mu_{u, b_{-i}}(B) du \neq 0$, or equivalently, the signed measure $\nu_{b,i}$ is absolutely continuous w.r.t. measure μ_b .
3. The payment functions $\lambda_i(A(\hat{b}), \hat{b}, b)$ satisfy

$$\lambda_i(A(\hat{b}), \hat{b}, b) = \rho_{b,i}^\mu(\hat{b}) A_i(\hat{b}) + \lambda_i^0(\hat{b}, b) \text{ a.s.}$$

where $\mathbf{E}_{\hat{b} \sim \mu_b}[\lambda_i^0(\hat{b}, b)] = 0$ and $\rho_{b,i}^\mu(\hat{b})$ is the density function relating $\nu_{b,i}$ to μ_b .

(Almost surely, or a.s., means that it holds everywhere except for a set with measure zero under $\mu_b(\cdot)$.)

Proof:

Necessity We first prove the necessity of the three conditions above. The first condition, that \mathcal{A} is monotone in expectation, follows directly from Archer-Tardos characterization of truthful allocation functions. The second and third conditions, as we prove below, are necessary for the expected payment to take the form required by the Archer-Tardos characterization.

We now write down the truthful payments give by the Archer-Tardos characterization, and rewrite it using the signed measure $\nu_{b,i}$.

$$\begin{aligned} \mathbf{E}[\mathcal{P}_i] &= b_i \mathbf{E}[\mathcal{A}_i(b)] - \int_0^{b_i} \mathbf{E}[\mathcal{A}_i(u, b_{-i})] du \\ &= b_i \mathbf{E}_{\hat{b} \sim \mu_b}[A_i(\hat{b})] - \int_0^{b_i} \mathbf{E}_{\hat{b} \sim \mu_{u, b_{-i}}}[A_i(\hat{b})] du \\ &= \int_{\hat{b} \in \mathbb{R}^n} A_i(\hat{b}) d\nu_{b,i} . \end{aligned}$$

where the last equality follows from the definition of the signed measure $\nu_{b,i}$, and Lemma C.1.

By definition of the reduction, we can write the expected payment as:

$$\mathbf{E}[\mathcal{P}_i] = \int_{\hat{b} \in \mathbb{R}^n} \lambda_i(A(\hat{b}), \hat{b}, b) d\mu_b .$$

Equating these two gives

$$\int_{\hat{b} \in \mathbb{R}^n} \lambda_i(A(\hat{b}), \hat{b}, b) d\mu_b = \mathbf{E}[\mathcal{P}_i] = \int_{\hat{b} \in \mathbb{R}^n} A_i(\hat{b}) d\nu_{b,i} . \quad (15)$$

Next, we define the normalized payment function $\tilde{\lambda}$ as

$$\tilde{\lambda}_i(A(\hat{b}), \hat{b}, b) = \lambda_i(A(\hat{b}), \hat{b}, b) - \lambda_i(0^n, \hat{b}, b) .$$

By (15), $\int_{\hat{b} \in \mathbb{R}^n} \lambda_i(0^n, \hat{b}, b) d\mu_b(B) = 0$, and therefore we may write

$$\int_{\hat{b} \in \mathbb{R}^n} \tilde{\lambda}_i(A(\hat{b}), \hat{b}, b) d\mu_b = \int_{\hat{b} \in \mathbb{R}^n} A_i(\hat{b}) d\nu_{b,i} .$$

If the above equality were to hold for all bounded, monotone, measurable allocation functions A , then by Lemma C.3, this implies for all Borel measurable sets $X \subseteq \mathbb{R}^n$:

$$\int_{\hat{b} \in X} \tilde{\lambda}_i(A(\hat{b}), \hat{b}, b) d\mu_b = \int_{\hat{b} \in X} A_i(\hat{b}) d\nu_{b,i} . \quad (16)$$

This statement would be intuitive if we allowed A_i to be any function — we could pick the function $A'_i(b) = 1_X(b)A_i(b)$, i.e. we could zero A_i except on X , and plug back into the previous equality. Unfortunately, this A'_i is not monotone. The work of Lemma C.3 is to show that the space of bounded, monotone functions is still sufficiently general as to guarantee equality for any Borel measurable set X .

Having derived Equation (16), we now show how it makes conditions two and three in theorem necessary. If we substitute the constant function $A_i(b) = 1$ into (16), we see that for all measurable X

$$\int_{\hat{b} \in X} \tilde{\lambda}_i(1^n, \hat{b}, b) d\mu_b = \int_{\hat{b} \in X} d\nu_{b,i} ,$$

that is, $\tilde{\lambda}_i(1^n, \hat{b}, b)$ satisfies the definition of the derivative of $\nu_{b,i}$ w.r.t μ_b , and therefore $\rho_{b,i}^\mu(\hat{b}) = \tilde{\lambda}_i(1^n, \hat{b}, b)$. Thus, given that finite payments λ exist it follows that the density relating $\nu_{b,i}$ to μ_b , namely $\rho_{b,i}^\mu(\hat{b})$, also exists and is finite. But given that both μ_b and $\nu_{b,i}$ are finite measures, this also means that $\nu_{b,i}$ is absolutely continuous w.r.t. μ_b . If not, then there exists a Borel measurable set V such that $\nu_{b,i}(V) \neq 0$ but $\mu_b(V) = 0$. We run into an immediate contradiction as follows:

$$0 = \int_{\hat{b} \in V} \rho_{b,i}^\mu(\hat{b}) d\mu_b = \int_{\hat{b} \in V} d\nu_{b,i} = \nu_{b,i}(V) \neq 0.$$

Thus we have proved that condition two, absolute continuity of $\nu_{b,i}$ w.r.t. μ_b , is necessary.

Returning to (16), by the definition of $\rho_{b,i}^\mu(\hat{b})$ we can write

$$\begin{aligned} \int_{\hat{b} \in X} \tilde{\lambda}_i(A(\hat{b}), \hat{b}, b) d\mu_b &= \int_{\hat{b} \in X} A_i(\hat{b}) \rho_{b,i}^\mu(\hat{b}) d\mu_b \\ \int_{\hat{b} \in X} \left(\tilde{\lambda}_i(A(\hat{b}), \hat{b}, b) - A_i(\hat{b}) \rho_{b,i}^\mu(\hat{b}) \right) d\mu_b &= 0 \end{aligned}$$

for all Borel measurable sets $X \subseteq \mathbb{R}^n$. By a standard argument (Fact E.12), this implies

$$\tilde{\lambda}_i(A(\hat{b}), \hat{b}, b) - A_i(\hat{b}) \rho_{b,i}^\mu(\hat{b}) = 0$$

almost surely with respect to $\mu_b(B)$, the third condition. Thus we have shown that all the three conditions are necessary.

Sufficiency We now show that the three stated conditions are sufficient. In a single-parameter setting, for a mechanism to be truthful, we simply need the allocation function to be monotone in expectation, and the payment function must satisfy the Archer-Tardos payment functions. Condition one guarantees that the allocation function output by the single-call reduction is a monotone in expectation allocation function. It remains to show that the second and third conditions result in payments that agree with Archer-Tardos payments. Given condition two, we see that $\nu_{b,i}$ is absolutely continuous w.r.t the resampling measure μ_b , and thus by Radon Nikodym theorem, the density function $\rho_{b,i}^\mu(\cdot)$ is finite and exists. All we need to show is that under the formula of $\lambda_i(A(\hat{b}), \hat{b}, b)$ described in condition three, we have

$$\int_{\hat{b} \in \mathbb{R}^n} \lambda_i(A(\hat{b}), \hat{b}, b) d\mu_b = b_i \mathbf{E}[\mathcal{A}_i(b)] - \int_0^{b_i} \mathbf{E}[\mathcal{A}_i(u, b_{-i})] du.$$

Once we substitute the formula for $\lambda_i(A(\hat{b}), \hat{b}, b)$ from condition three, this equality follows from the definition of $\rho_{b,i}^\mu(\cdot)$ and $\nu_{b,i}$. \blacksquare

Lemma C.3 *Let μ and ν be finite measures (possibly signed), and let $g : \mathbb{R}_+^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ be a function with $g(0, \hat{b}) = 0$ satisfying*

$$\int_{\hat{b} \in \mathbb{R}^n} g(A(\hat{b}), \hat{b}) d\mu = \int_{\hat{b} \in \mathbb{R}^n} A_i(\hat{b}) d\nu$$

for all Borel measurable functions $A : \mathbb{R}^n \rightarrow \mathbb{R}_+^n$ where A is bounded and monotone in the sense that $b' \geq b \Rightarrow A(b') \geq A(b)$.

Then for any such A and all Borel measurable sets $X \subseteq \mathbb{R}^n$,

$$\int_{\hat{b} \in X} g(A(\hat{b}), \hat{b}) d\mu = \int_{\hat{b} \in X} A_i(\hat{b}) d\nu .$$

Proof: First, assume that the characteristic function of X can be written as the difference of two $\{0, 1\}$ monotone functions, that is, $1_X(b) = f^+(b) - f^-(b)$ where f^+ and f^- are monotone functions mapping \mathbb{R}^n to $\{0, 1\}$. Note that this includes all rectangular parallelepipeds (a product of open, closed, or half-open intervals).

Define as $A_i^+(b) = A_i(b) \cdot f^+(b)$ and $A_i^-(b) = A_i(b) \cdot f^-(b)$. Note that for any bounded, monotone, measurable A , the functions A^+ and A^- are similarly bounded and monotone. Therefore the conditions of the lemma imply

$$\int_{\hat{b} \in \mathbb{R}^n} g(A^+(\hat{b}), \hat{b}) d\mu = \int_{\hat{b} \in \mathbb{R}^n} A_i^+(\hat{b}) d\nu$$

and

$$\int_{\hat{b} \in \mathbb{R}^n} g(A^-(\hat{b}), \hat{b}) d\mu = \int_{\hat{b} \in \mathbb{R}^n} A_i^-(\hat{b}) d\nu$$

Taking the difference, we get

$$\int_{\hat{b} \in \mathbb{R}^n} \left(g(A^+(\hat{b}), \hat{b}) - g(A^-(\hat{b}), \hat{b}) \right) d\mu = \int_{\hat{b} \in \mathbb{R}^n} \left(A_i^+(\hat{b}) - A_i^-(\hat{b}) \right) d\nu .$$

Note that $A^+ = A^-$ everywhere except on the set X , so the integrands are only nonzero on X , thus we can replace \mathbb{R}^n with X in the integrals:

$$\int_{\hat{b} \in X} \left(g(A^+(\hat{b}), \hat{b}) - g(A^-(\hat{b}), \hat{b}) \right) d\mu = \int_{\hat{b} \in X} \left(A_i^+(\hat{b}) - A_i^-(\hat{b}) \right) d\nu .$$

Now, note that on X , $A^+ = A$ and $A^- = 0$. Thus, also using the fact $g(A^-(\hat{b}), \hat{b}) = 0$, we have

$$\int_{\hat{b} \in X} g(A(\hat{b}), \hat{b}) d\mu = \int_{\hat{b} \in X} A_i(\hat{b}) d\nu ,$$

as desired.

To show that the lemma holds for all Borel measurable sets X , we observe that it holds for all rectangular parallelepipeds (a product of open, closed, or half-open intervals) by the above argument. Since the set of rectangular parallelepipeds is closed under finite intersections, the lemma applies to all finite intersections of rectangular parallelepipeds, which is the π -system that generates the Borel σ -algebra of \mathbb{R}^n .

Additionally, if the lemma holds for a countable sequence of disjoint sets X_k , then it clearly holds for their union as well, implying that the sets for which the lemma is true must be a λ -system.

Therefore, by Dynkin's π - λ theorem, the λ -system (the sets satisfying the lemma) must contain all sets in the σ -algebra generated by the π -system (the set of rectangular parallelepipeds) — namely, it must contain all sets in the Borel σ -algebra of \mathbb{R}^n . Thus, the lemma must hold for all Borel measurable sets X . ■

C.1 Monotonicity and μ_b

Theorem C.2 requires μ_b to be such that $\mathcal{A}_i(b)$ is monotone in expectation. The following lemma gives a necessary condition:

Lemma C.4 *Let B be a set of bids that is leftward closed with respect to b_i , i.e. if $\hat{b} \in B$, then $(u, \hat{b}_{-i}) \in B$ for all $u \in (-\infty, b_i] \cap T_i$. If $\mu_b(B)$ satisfies the monotonicity condition*

$$\Pr(\hat{b} \in B | b) = \mu_b(B)$$

is weakly decreasing in b_i . Similarly, if B is rightward closed with respect to b_i (i.e. $\hat{b} \in B$ implies $\hat{b}_{-i}u \in B$ for $u \in [b_i, \infty)$), then $\Pr(\hat{b} \in B | b)$ is weakly increasing in b_i , and if B is both rightward and leftward closed with respect to b_i then $\Pr(\hat{b} \in B | b)$ is constant in b_i .

Proof: First, we prove the case where B rightward closed. For contradiction, let B be a rightward closed set on which f violates the statement of the lemma for some b and $b'_i > b_i$, i.e.

$$\Pr(\hat{b} \in B | b) = \mu_b(B) > \mu_{b'_i, b_{-i}}(B) = \Pr(\hat{b} \in B | b'_i, b_{-i}) .$$

Consider the monotone allocation function

$$A_i(b) = \begin{cases} 1, & b \in B \\ 0, & \text{otherwise.} \end{cases}$$

Noting that the $\mathbb{E}[\mathcal{A}(b)] = \mu_b(B)$, we have

$$E[\mathcal{A}_i(b'_i, b_{-i})] = \mu_{b'_i, b_{-i}}(B) < \mu_b(B) = E[\mathcal{A}_i(b)] .$$

Thus, under this allocation function, bidder i lowers her expected utility by raising her bid to b'_i , contradicting the monotonicity condition.

Finally, any leftward closed set B is the complement (probabilistically) of a rightward closed set, therefore $\Pr(\hat{b} \in B | b)$ must be weakly decreasing. For a set B that is both leftward and rightward closed, the theorem follows because $\Pr(\hat{b} \in B | b)$ must be both weakly increasing and weakly decreasing. ■

D Optimality proofs for generalized BKS

In this section, we generalize our optimality result of Section 6.3 to arbitrary probability measures and give a complete proof. Theorem C.2 shows that truthful payments take the form

$$\lambda_i(A(\hat{b}), \hat{b}, b) = \rho_b^\mu(\hat{b}) A_i(\hat{b}) + \lambda_i^0(\hat{b}, b) \text{ a.s.}$$

and thus optimizing the bid-normalized payments means optimizing the following quantity:

$$\sum_j \frac{\lambda_{ij}(b_j(A(\hat{b})), \hat{b}, b)}{b_j(A(\hat{b}))} = \frac{\rho_b^\mu(\hat{b}) A_i(\hat{b})}{b_i A_i(\hat{b})} = \frac{\rho_b^\mu(\hat{b})}{b_i}.$$

This means that for worst-case payments we will optimize $\sup_{i, \hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right|$, and for payment variance we will optimize $\max_i \text{Var}_{\hat{b} \sim \mu_b} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right)$. We show that the BKS transformation is optimal for both, subject to an almost everywhere caveat:

Theorem D.1 (Optimality of the BKS Transformation) *(Generalization of Theorem 6.3) The BKS reduction $\text{SPtoMechBKS}(A, \gamma)$ optimizes the payment variance and worst-case normalized payment subject to a lower bound of $\alpha = (1 - \gamma)^n \in (\frac{1}{e}, 1)$ on the precision, the welfare approximation ($n \geq 2$), or the revenue approximation ($n \geq 2$). That is, for any other truthful reduction $(\mu, \{\lambda_i\})$ that achieves a precision, welfare approximation, or revenue approximation of α , the worst-case normalized payments are at least as large almost everywhere over b :*

$$\sup_{A, i} \text{Var}_{\hat{b} \sim \mu_b} \left(\sum_j \frac{\lambda_{ij}(b_j(A(\hat{b})), \hat{b}, b)}{b_j(A(\hat{b}))} \right) = \max_i \text{Var}_{\hat{b} \sim \mu_b} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right) \geq \max_i \text{Var}_{\hat{b} \sim \mu_b} \left(\frac{\rho_b^{\text{BKS}}(\hat{b})}{b_i} \right) \text{ a.e.}$$

and

$$\sup_{A, i, \hat{b}} \left| \sum_j \frac{\lambda_{ij}(b_j(A(\hat{b})), \hat{b}, b)}{b_j(A(\hat{b}))} \right| = \sup_{i, \hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \sup_{i, \hat{b}} \left| \frac{\rho_b^{\text{BKS}}(\hat{b})}{b_i} \right| \text{ a.e.}$$

Under the nice distribution assumption, this holds for every b .

The theorem is proven in two steps. First, we prove in Theorem D.2 that the BKS transform optimizes precision. Second, we show in Lemma D.3 that a distribution which optimizes precision also optimizes the welfare and revenue approximations.

Theorem D.2 (Precision Optimality of the BKS Transformation) *(Generalization of Theorem 6.3) The BKS reduction $\text{SPtoMechBKS}(A, \gamma)$ optimizes the variance of normalized payments and the worst-case normalized payment subject to a lower bound of $\alpha_P = (1 - \gamma)^n \in (\frac{1}{e}, 1)$ on the precision almost everywhere over b . Under the nice distribution assumption, it is optimal for every b .*

Theorem D.2 is given in Sections D.1-D.3. Section D.1 defines probabilities that are used in the proof. Section D.2 proves Theorem D.2 with forward references to two important technical lemmas given in Section D.3.

Lemma D.3 *(Generalization of Lemma 6.4) For $\alpha > \frac{1}{e}$ and $n \geq 2$, a probability measure that optimizes the variance of normalized payments or the maximum normalized payment subject to a precision constraint of $\Pr(\hat{b} = b|b) \geq \alpha$ also optimizes the maximum normalized payment almost everywhere subject to a welfare or revenue approximation of α .*

Lemma D.3 is proven in Section D.4, building on technical lemmas from Section D.3.

D.1 Definitions

To prove Theorem D.2, we give names to certain probabilities. As in the MIDR setting, we use a set $M \subseteq [n]$ to denote the set of bidders with $\hat{b}_i = b_i$. Bidders $i \notin M$ have their bids lowered, that is $\hat{b}_i < b_i$. We define the probability $\pi^\mu(M, b)$ to be the probability that such an event occurs, that is, $\pi^\mu(M, b)$ is the probability when b is bid that $\hat{b}_i = b_i$ if $i \in M$, and $\hat{b}_i < b_i$ if $i \notin M$:

$$\pi^\mu(M, b) \equiv \Pr \left((\hat{b}_i = b_i \text{ for } i \in M) \text{ and } (\hat{b}_i < b_i \text{ for } i \notin M) \middle| b \right).$$

Note that for the BKS transformation, $\pi^\mu(M, b) = (1 - \gamma)^{|M|} \gamma^{n-|M|}$ so $\frac{\pi^\mu(M \cup \{i\}, b)}{\pi^\mu(M, b)} = \frac{1-\gamma}{\gamma}$.

The second probability quantifies the behavior of μ_b near b as follows. Fix a bid b and assume player i actually bids $b_i - \delta$. Does the distribution $\mu_{b_i - \delta, b_{-i}}$ cause the reduction to select $\hat{b} = b$ with positive probability in spite of the fact that i said $b_i - \delta$? In particular, we care about the average behavior for $\delta \in [0, b_i]$, which we represent by $z^\mu(M, i, b)$. Formally, we define

$$\zeta^\mu(M, i, b, z) \equiv \Pr \left(\hat{b}_i = z \text{ and } (\hat{b}_j = b_j \text{ for } j \in M \setminus \{i\}) \text{ and } (\hat{b}_j < b_j \text{ for } j \notin M \cup \{i\}) \middle| b \right)$$

and

$$z^\mu(M, i, b) \equiv \frac{1}{b_i} \int_0^{b_i} \zeta^\mu(M, i, (u, b_{-i}), b_i) \, du.$$

Of particular importance, we will show $z^\mu(M, i, b) = 0$ almost everywhere in general and everywhere under the nice distribution assumption.

D.2 Precision Optimality of the BKS Transformation

The optimality proof for the BKS transformation

The first result follows as a corollary of Lemma D.8:

Corollary D.4 (of Lemma D.8) *If a resampling distribution μ satisfies the monotonicity condition, then for all $M, i \notin M$:*

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{\pi^\mu(M \cup \{i\}, b) - z^\mu(M, i, b)}{\pi^\mu(M, b)}$$

and

$$\begin{aligned} & \int_{\hat{b}_i \leq b_i \wedge (j \in M \Rightarrow \hat{b}_j = b_j) \wedge (j \notin M \cup \{i\} \Rightarrow \hat{b}_j < b_j)} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right)^2 d\mu_b \\ & \geq (\pi^\mu(M, b) + \pi^\mu(M \cup \{i\}, b)) \frac{\pi^\mu(M \cup \{i\}, b)}{\pi^\mu(M, b)} \left(1 - \frac{z^\mu(M, i, b)}{\pi^\mu(M \cup \{i\}, b)} \right)^2. \end{aligned}$$

Proof: Apply Lemma D.8 where B_{-i} is the set of \hat{b}_{-i} where $\hat{b}_j = b_j$ if $j \in M$ and $\hat{b}_j < b_j$ for $j \notin M$. ■

If we ignore the $z^\mu(M, i, b)$ terms, this looks precisely like the normalized payments from the MIDR setting. Fortunately, $z^\mu(M, i, b)$ is almost always zero:

Corollary D.5 (of Lemma D.11) *For any resampling distribution μ and a fixed M and i ,*

$$z^\mu(M, i, b) = 0 \text{ a.e.}$$

(i.e. for all but a set of b with zero measure).

Under the nice distribution assumption, $z^\mu(M, i, b) = 0$ for all b .

Proof: Note that $\zeta^\mu(M, i, (u, b_{-i}), b_i) \leq \Pr_\mu(\hat{b}_i = b_i | u, b_{-i})$, so by Lemma D.11

$$z^\mu(M, i, b) = \frac{1}{b_i} \int_0^{b_i} \zeta^\mu(M, i, (u, b_{-i}), b_i) \leq \int_0^{b_i} \Pr_\mu(\hat{b}_i = b_i | u, b_{-i}) = 0 \text{ a.e.}$$

■

Thus, Corollaries D.4 and D.5 together imply the following bound:

Lemma D.6 *If a resampling distribution μ with precision $\alpha \geq (1-\gamma)^n$ satisfies the monotonicity condition, then*

$$\sup_{i, \hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{1-\gamma}{\gamma} \text{ a.e.}$$

that is, for all b but a set with measure zero. This holds everywhere if $z^\mu(M, i, b) = 0$ everywhere.

Proof: We first prove the bound on the worst-case normalized payment. By assumption on the precision of μ , we have $\pi^\mu([n], b) \geq (1-\gamma)^n$ for some γ and all b . By Corollary D.4, we know that

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{\pi^\mu(M \cup \{i\}, b) - z^\mu(M, i, b)}{\pi^\mu(M, b)}.$$

Applying Lemma D.9 with $\eta(S) = \pi^\mu(S, b)$, $\alpha = (1-\gamma)^n$, and $\beta = 1$ we get that

$$\max_{M, i \notin M} \frac{\pi^\mu(M \cup \{i\}, b)}{\pi^\mu(M, b)} \geq \frac{1-\phi}{\phi}$$

where

$$\phi = 1 - \left(\frac{(1-\gamma)^n}{1} \right)^{\frac{1}{n}} = \gamma.$$

Thus,

$$\max_{M, i \notin M} \frac{\pi^\mu(M \cup \{i\}, b)}{\pi^\mu(M, b)} \geq \frac{1-\gamma}{\gamma}.$$

Aggregating Corollary D.4 over all M and $i \notin M$, we have

$$\begin{aligned} \sup_{i, \hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| &\geq \max_{M, i \notin M} \frac{\pi^\mu(M \cup \{i\}, b)}{\pi^\mu(M, b)} - \max_{M, i \notin M} \frac{z^\mu(M, i, b)}{\pi^\mu(M, b)} \\ &\geq \frac{1-\gamma}{\gamma} - \max_{M, i \notin M} \frac{z^\mu(M, i, b)}{\pi^\mu(M, b)}. \end{aligned}$$

If we assume $z^\mu(M, i, b) = 0$ everywhere (e.g. by the nice distribution assumption), then we get

$$\sup_{i, \hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{1-\gamma}{\gamma}.$$

Otherwise, Corollary D.5 says that $z^\mu(M, i, b) = 0$ almost everywhere, giving the more general bound. ■

Lemma D.7 *If a resampling distribution μ with precision $\alpha \geq (1 - \gamma)^n \geq \frac{1}{e}$ satisfies the monotonicity condition, then*

$$\max_i \text{Var}_{\hat{b}} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right) \geq \frac{1 - \gamma}{\gamma} \text{ a.e.}$$

that is, for all b but a set with measure zero. This holds everywhere if $z^\mu(M, i, b) = 0$ everywhere.

Proof: The proof for variance is similar to Lemma D.6, but we apply Lemma D.10 instead of Lemma D.9. First, note that since μ_b is a probability measure, $\mu_b(\mathbb{R}^n) = 1$ and thus

$$\int_{\hat{b} \in \mathbb{R}^n} \frac{\rho_b^\mu(\hat{b})}{b_i} d\mu_b = \frac{1}{b_i} \nu_{b,i}(\mathbb{R}^n) = \mu_b(\mathbb{R}^n) - \frac{1}{b_i} \int_0^{b_i} \mu_{u, b-i}(\mathbb{R}^n) du = 1 - \frac{1}{b_i} \int_0^{b_i} 1 du = 0 \text{ .}$$

We begin with the variance for player i , applying Corollaries D.4 and D.5:

$$\begin{aligned} \text{Var}_{\hat{b}} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right) &= \int_{\hat{b} \in \mathbb{R}^n} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right)^2 d\mu_b - \left(\int_{\hat{b} \in \mathbb{R}^n} \frac{\rho_b^\mu(\hat{b})}{b_i} d\mu_b \right)^2 \\ &= \int_{\hat{b} \in \mathbb{R}^n} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right)^2 d\mu_b \\ &\geq \sum_{M|i \notin M} \int_{\hat{b}_i \leq b_i \wedge (j \in M \Rightarrow \hat{b}_j = b_j) \wedge (j \notin M \cup \{i\} \Rightarrow \hat{b}_j < b_j)} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right)^2 d\mu_b \\ &\geq \sum_{M|i \notin M} (\pi^\mu(M, b) + \pi^\mu(M \cup \{i\}, b)) \frac{\pi^\mu(M \cup \{i\}, b)}{\pi^\mu(M, b)} \text{ a.e.} \end{aligned}$$

Applying Lemma D.10 with $\eta(S) = \pi^\mu(S, b)$, $\alpha = (1 - \gamma)^n$ and $\beta = \Pr(\hat{b} \leq b | b)$ immediately implies

$$\max_i \text{Var}_{\hat{b}} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right) \geq \Pr(\hat{b} \leq b | b) \frac{1 - \phi}{\phi} \text{ a.e.}$$

where

$$\phi = 1 - \left(\frac{(1 - \gamma)^n}{\Pr(\hat{b} \leq b | b)} \right)^{\frac{1}{n}} \text{ .}$$

One can check that when $\frac{(1 - \gamma)^n}{\Pr(\hat{b} \leq b | b)} \geq \frac{1}{e}$, the quantity $\Pr(\hat{b} \leq b | b) \frac{1 - \phi}{\phi}$ is decreasing in $\Pr(\hat{b} \leq b | b)$.

Taking the worst case $\Pr(\hat{b} \leq b | b) = 1$ implies the desired result:

$$\max_i \text{Var}_{\hat{b}} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right) \geq \max_i \geq \frac{1 - \gamma}{\gamma} \text{ a.e.}$$

■

Theorem D.2 – optimality of the BKS transformation with respect to a precision bound – follows from the two previous lemmas:

Proof:[of Theorem D.2] For worst-case payments, we show that for any measure μ , with precision at least 2^{-n} ,

$$\sup_{i, \hat{b}} \left| \frac{\rho_b^{BKS}(\hat{b})}{b_i} \right| \leq \sup_{i, \hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \quad a.e.$$

For $\Pr(\hat{b} = b|b) = (1 - \gamma)^n$, the BKS transform achieves $\sup_{i, \hat{b}} \left| \frac{\rho_b^{BKS}(\hat{b})}{b_i} \right| = \max\left(1, \frac{1-\gamma}{\gamma}\right)$ for all b . Provided $\gamma > \frac{1}{2}$, the dominant term is $\frac{1-\gamma}{\gamma}$ and Lemma D.6 shows that this is a lower bound for any such μ almost everywhere. When $\alpha > 2^{-n}$ we get $\gamma > \frac{1}{2}$, and thus BKS is optimal.

Moreover, under the nice distribution assumption (implying $z^\mu(M, i, b) = 0$), Lemma D.6 says that this holds everywhere.

For the variance of normalized payments, we need to show that for any measure μ with precision at least $\frac{1}{e}$:

$$\text{Var}_{\hat{b} \sim \mu_b} \left(\frac{\rho_b^{BKS}(\hat{b})}{b_i} \right) \leq \text{Var}_{\hat{b} \sim \mu_b} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right) \quad a.e.$$

Again, for $\Pr(\hat{b} = b|b) = (1 - \gamma)^n$, the BKS transform achieves $\text{Var}_{\hat{b} \sim \mu_b} \left| \frac{\rho_b^{BKS}(\hat{b})}{b_i} \right| = \frac{1-\gamma}{\gamma}$ for all b . Lemma D.7 shows that this is a lower bound for any such μ almost everywhere. ■

D.3 Technical Lemmas

The next lemma gives our main lower bound on the worst coefficient:

Lemma D.8 *If a measure μ satisfies the monotonicity condition, then for any player i , bid b , and set of bids $B_{-i} \subseteq \mathbb{R}_+^{n-1}$:*

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{\Pr(\hat{b}_i = b_i \wedge \hat{b}_{-i} \in B_{-i} | b) - \frac{1}{b_i} \int_0^{b_i} \Pr(\hat{b}_i = b_i \wedge \hat{b}_{-i} \in B_{-i} | u, b_{-i})}{\Pr(\hat{b}_i < b_i \wedge \hat{b}_{-i} \in B_{-i} | b)},$$

and

$$\begin{aligned} \int_{\hat{b}_i \leq b_i \wedge \hat{b}_{-i} \in B_{-i}} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right)^2 d\mu_b &\geq \Pr(\hat{b}_i \leq b_i \wedge \hat{b}_{-i} \in B_{-i} | b) \frac{\Pr(\hat{b}_i = b_i \wedge \hat{b}_{-i} \in B_{-i} | b)}{\Pr(\hat{b}_i < b_i \wedge \hat{b}_{-i} \in B_{-i} | b)} \\ &\times \left(1 - \frac{\frac{1}{b_i} \int_0^{b_i} \Pr(\hat{b}_i = b_i \wedge \hat{b}_{-i} \in B_{-i} | u, b_{-i})}{\Pr(\hat{b}_i = b_i \wedge \hat{b}_{-i} \in B_{-i} | b)} \right)^2, \end{aligned}$$

where the integral terms are zero almost everywhere in b by Lemma D.11.

Proof: Define the sets

$$B^{(=)} = \{b_i\} \times B_{-i} \quad \text{and} \quad B^{(<)} = [0, b_i) \times B_{-i},$$

i.e. the set $B^{(=)}$ contains bids \hat{b} where $\hat{b}_i = b_i$ and $\hat{b}_{-i} \in B_{-i}$, and the set $B^{(<)}$ contains bids \hat{b} where $\hat{b}_i < b_i$ and $\hat{b}_{-i} \in B_{-i}$. The main work of the lemma is to bound the following term:

$$\begin{aligned}
\int_{\hat{b} \in B^{(<)}} \frac{\rho_b^\mu(\hat{b})}{b_i} d\mu_b &= \frac{\nu_{b,i}(B^{(<)})}{b_i} \\
&= \frac{b_i \mu_b(B^{(<)}) - \int_0^{b_i} \mu_{u,b_{-i}}(B^{(<)}) du}{b_i} \\
&= \mu_b(B^{(<)}) - \frac{1}{b_i} \int_0^{b_i} \mu_{u,b_{-i}}(B^{(<)}) du \\
&= \Pr(\hat{b} \in B^{(<)} | b) - \frac{1}{b_i} \int_0^{b_i} \Pr(\hat{b} \in B^{(<)} | u, b_{-i}) du .
\end{aligned}$$

By monotonicity, $\Pr(\hat{b} \in B^{(<) \cup (=)} | u, b_{-i})$ is weakly decreasing in u (Lemma C.4). This implies

$$\Pr(\hat{b} \in B^{(=)} | b) + \Pr(\hat{b} \in B^{(<)} | b) \leq \frac{1}{b_i} \int_0^{b_i} \left(\Pr(\hat{b} \in B^{(=)} | u, b_{-i}) + \Pr(\hat{b} \in B^{(<)} | u, b_{-i}) \right) du$$

and thus

$$\begin{aligned}
\int_{\hat{b} \in B^{(<)}} \frac{\rho_b^\mu(\hat{b})}{b_i} d\mu_b &= \Pr(\hat{b} \in B^{(<)} | b) - \frac{1}{b_i} \int_0^{b_i} \Pr(\hat{b} \in B^{(<)} | u, b_{-i}) du \\
&\leq - \left(\Pr(\hat{b} \in B^{(=)} | b) - \frac{1}{b_i} \int_0^{b_i} \Pr(\hat{b} \in B^{(=)} | u, b_{-i}) du \right) .
\end{aligned}$$

To bound $\sup_{\hat{b} \in B^{(<)}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right|$, we have

$$\sup_{\hat{b} \in B^{(<)}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \left| \frac{\int_{\hat{b} \in B^{(<)}} \frac{\rho_b^\mu(\hat{b})}{b_i} d\mu_b}{\mu_b(B^{(<)})} \right| \geq \frac{\Pr(\hat{b} \in B^{(=)} | b) - \frac{1}{b_i} \int_0^{b_i} \Pr(\hat{b} \in B^{(=)} | u, b_{-i}) du}{\Pr(\hat{b} \in B^{(<)} | b)}$$

Lemma D.11 implies that the limit term is zero almost everywhere in b .

For our partial bound on the second moment, we write

$$\begin{aligned}
\int_{\hat{b} \in B^{(<)} \cup B^{(=)}} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right)^2 d\mu_b &\geq \int_{\hat{b} \in B^{(=)}} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right)^2 d\mu_b + \int_{\hat{b} \in B^{(<)}} \left(\frac{\rho_b^\mu(\hat{b})}{b_i} \right)^2 d\mu_b \\
&\geq \mu_b(B^{(=)}) \left(\frac{\int_{\hat{b} \in B^{(=)}} \frac{\rho_b^\mu(\hat{b})}{b_i} d\mu_b}{\mu_b(B^{(=)})} \right)^2 + \mu_b(B^{(<)}) \left(\frac{\int_{\hat{b} \in B^{(<)}} \frac{\rho_b^\mu(\hat{b})}{b_i} d\mu_b}{\mu_b(B^{(<)})} \right)^2 \\
&\geq \mu_b(B^{(=)}) \left(\frac{\Pr(\hat{b} \in B^{(=)} | b) - \frac{1}{b_i} \int_0^{b_i} \Pr(\hat{b} \in B^{(=)} | u, b_{-i}) du}{\mu_b(B^{(=)})} \right)^2 \\
&\quad + \mu_b(B^{(<)}) \left(\frac{\Pr(\hat{b} \in B^{(=)} | b) - \frac{1}{b_i} \int_0^{b_i} \Pr(\hat{b} \in B^{(=)} | u, b_{-i}) du}{\mu_b(B^{(<)})} \right)^2 \\
&\geq \left(\mu_b(B^{(<)}) + \mu_b(B^{(=)}) \right) \frac{\mu_b(B^{(=)})}{\mu_b(B^{(<)})} \\
&\quad \times \left(1 - \frac{\frac{1}{b_i} \int_0^{b_i} \Pr(\hat{b} \in B^{(=)} | u, b_{-i}) du}{\Pr(\hat{b} \in B^{(=)} | b)} \right)^2
\end{aligned}$$

Which is the desired bound. ■

Lemma D.9 Let $\eta : \{0, 1\}^n$ be a function over subsets $S \subseteq [n]$ with $\eta([n]) \geq \alpha \in [0, 1]$ and $\sum_{S \subseteq [n]} \eta(S) \leq \beta \in [0, 1]$. Then

$$\max_{S, i \in [n] \setminus S} \frac{\eta(S \cup \{i\})}{\eta(S)} \geq \frac{1 - \phi}{\phi}$$

where $\phi = 1 - \left(\frac{\alpha}{\beta} \right)^{\frac{1}{n}}$.

Proof: By contradiction. Assume that for every S and $i \notin S$,

$$\frac{\eta(S \cup \{i\})}{\eta(S)} < \frac{1 - \phi}{\phi}$$

where $\phi = 1 - \left(\frac{\alpha}{\beta} \right)^{\frac{1}{n}}$.

Then by multiplying $\frac{\eta(S)}{\eta(S \cup \{i\})}$ terms together we get

$$\eta(S) \geq \eta([n]) \left(\frac{\phi}{1 - \phi} \right)^{n - |S|}.$$

Summing over all $S \subseteq [n]$, substituting for α and β , and algebra gives

$$\begin{aligned}
\sum_{S \subseteq [n]} \eta(S) &> \eta([n]) \sum_{S \subseteq [n]} \left(\frac{\phi}{1-\phi} \right)^{n-|S|} \\
\beta &> \alpha \sum_{S \subseteq [n]} \left(\frac{\phi}{1-\phi} \right)^{n-|S|} \\
\beta(1-\phi)^n &> \alpha \sum_{S \subseteq [n]} (1-\phi)^{|S|} \phi^{n-|S|} \\
\beta(1-\phi)^n &> \alpha \\
\beta \left(1 - \left(1 - \left(\frac{\alpha}{\beta} \right)^{\frac{1}{n}} \right) \right)^n &> \alpha \\
\alpha &> \alpha .
\end{aligned}$$

Which is a contradiction. ■

Lemma D.10 *Let $\eta : \{0, 1\}^n$ be a function over subsets $S \subseteq [n]$ with $\eta([n]) \geq \alpha \in [0, 1]$ and $\sum_{S \subseteq [n]} \eta(S) = \beta \in [0, 1]$. Then*

$$\max_i \sum_{S | i \notin S} (\eta(S) + \eta(S \cup \{i\})) \frac{\eta(S \cup \{i\})}{\eta(S)} \geq \beta \frac{1-\phi}{\phi}$$

where $\phi = 1 - \left(\frac{\alpha}{\beta} \right)^{\frac{1}{n}}$.

Proof: We lower-bound the sum. Fix i and differentiate the sum:

$$\frac{\partial}{\partial \eta(S)} \left(\sum_{T | i \notin T} (\eta(T) + \eta(T \cup \{i\})) \frac{\eta(T \cup \{i\})}{\eta(T)} \right) = \begin{cases} 2 \frac{\eta(S)}{\eta(S \setminus \{i\})} + 1, & i \in S \\ - \left(\frac{\eta(S \cup \{i\})}{\eta(S)} \right)^2, & i \notin S \end{cases}.$$

The conditions of the lemma bound $\sum_S \eta(S)$ and $\eta([n])$, otherwise the values of η are only constrained to be in $[0, 1]$. The derivative tells us that in an optimal assignment, for all sets S that do not contain i , the ratio $\frac{\eta(S \cup \{i\})}{\eta(S)}$ is constant. Construct such an optimal assignment and define ϕ_i as satisfying

$$\frac{\eta(S \cup \{i\})}{\eta(S)} = \frac{1-\phi_i}{\phi_i}$$

for all S that do not contain i . Note that this implies

$$\sum_{S | i \notin S} (\eta(S) + \eta(S \cup \{i\})) \frac{\eta(S \cup \{i\})}{\eta(S)} \geq \beta \frac{1-\phi_i}{\phi_i}$$

For any set S it follows that

$$\begin{aligned}
\eta(S) &= \eta([n]) \prod_{i \notin S} \frac{\phi_i}{1 - \phi_i} \\
\sum_{S \subseteq [n]} \eta(S) &= \eta([n]) \sum_{S \subseteq [n]} \prod_{i \notin S} \frac{\phi_i}{1 - \phi_i} \\
\beta \prod_{i \in [n]} (1 - \phi_i) &\geq \alpha \sum_{S \subseteq [n]} \prod_{i \in S} (1 - \phi_i) \prod_{i \notin S} \phi_i \\
\prod_{i \in [n]} (1 - \phi_i) &\geq \frac{\alpha}{\beta} .
\end{aligned}$$

This implies there is some i such that $\phi_i \leq 1 - \left(\frac{\alpha}{\beta}\right)^{\frac{1}{n}}$, which implies the lemma. \blacksquare

The next lemma is our main analysis lemma. We will ultimately use it to claim that our lower bound must hold almost everywhere for any μ :

Lemma D.11 *For any resampling distribution μ that satisfies the monotonicity condition, any bid b , and any bidder i ,*

$$\int_0^{b_i} \Pr_{\mu}(\hat{b}_i = b_i | u, b_{-i}) = 0 \quad a.e.$$

(i.e. for all but a set of b with zero measure).

Proof: Define the marginalized measure μ_b^i for a set of bids $B \subseteq \mathbb{R}$ as

$$\mu_b^i(B) \equiv \mu_b(\{b \in \mathbb{R}^n | b_i \in B\}) .$$

Note that

$$\mu_{u, b_{-i}}^i(\{b_i\}) = \Pr_{\mu}(\hat{b}_i = b_i | u, b_{-i})$$

and therefore our task is to show that

$$\lim_{u \rightarrow -b_i} \mu_{u, b_{-i}}^i(\{b_i\}) = 0 \quad a.e.$$

Next we show that for any b we can prove the desired limit is zero by proving that a related integral is zero. Assume that for some b we have

$$\lim_{u \rightarrow -b_i} \mu_{u, b_{-i}}^i(\{b_i\}) > 0 .$$

Then there exists a δ_b such that

$$\forall u \in (b_i - \delta_b, b_i) : \quad \mu_{u, b_{-i}}^i(\{b_i\}) > 0 .$$

Since $\mu_{u, b_{-i}}^i(\{b_i\})$ is nonnegative, this implies

$$\int_{u \in \mathbb{R}} \mu_{u, b_{-i}}^i(\{b_i\}) du \geq \int_{u \in (b_i - \delta_b, b_i)} \mu_{u, b_{-i}}^i(\{b_i\}) du > 0 .$$

Taking the contrapositive, it follows that if the integral is zero at a bid b then the limit is also zero:

$$\int_{u \in \mathbb{R}} \mu_{u, b_{-i}}^i(\{b_i\}) du = 0 \Rightarrow \lim_{u \rightarrow -b_i} \mu_{u, b_{-i}}^i(\{b_i\}) = 0 . \quad (17)$$

Henceforth, we will prove that $\int_{u \in \mathbb{R}} \mu_{u, b_{-i}}^i(\{b_i\}) du = 0$ almost everywhere.

We start with the integral

$$\int_{b \in \mathbb{R}} \int_{u \in \mathbb{R}} \mu_{u, b_{-i}}^i(\{b_i\}) du db .$$

Manipulating the integral and noting that $\int_{u \in \mathbb{R}} 1_{\{u\}}(\hat{b}_i) du = 0$, we get

$$\begin{aligned} \int_{b \in \mathbb{R}} \int_{u \in \mathbb{R}} \mu_{u, b_{-i}}^i(\{b_i\}) du db &= \int_{b \in \mathbb{R}} \int_{u \in \mathbb{R}} \mu_{b_i, b_{-i}}^i(\{u\}) du db \\ &= \int_{b \in \mathbb{R}^n} \int_{u \in \mathbb{R}} \int_{\hat{b}_i \in \mathbb{R}} 1_{\{u\}}(\hat{b}_i) d\mu_b^i du db \\ &= \int_{b \in \mathbb{R}^n} \int_{\hat{b}_i \in \mathbb{R}} \int_{u \in \mathbb{R}} 1_{\{u\}}(\hat{b}_i) du d\mu_b^i db \\ &= \int_{b \in \mathbb{R}^n} \int_{\hat{b}_i \in \mathbb{R}} 0 d\mu_b^i db \\ &= 0 \end{aligned}$$

(where integral rearrangements may be justified by Tonelli's Theorem). By Fact E.11, this implies

$$\int_{u \in \mathbb{R}} \mu_{u, b_{-i}}^i(\{b_i\}) du = 0 \text{ almost everywhere over } b,$$

which implies the desired result. ■

D.4 Welfare and Revenue Optimality

Under mild assumptions, one can show that optimizing precision is equivalent to optimizing the social welfare approximation or the revenue approximation. We include only the worst-case optimality proofs; the variance proof is similar, applying ideas from Lemma D.7.

The optimality proof is divided into two steps:

1. *Lemmas D.12 and D.13:* Show that the welfare/revenue approximation of a resampling distribution μ is essentially

$$\inf_b \min_{i \in [n]} \Pr \left(\hat{b}_i \geq b_i \text{ and } \hat{b}_{-i} = b_{-i} \mid b \right) .$$

The welfare and revenue lemmas use different techniques to give a lower bound on the approximation; however, they use the same “bad” allocation function.

2. *Lemma D.14 and finally Lemma D.3:* Show that a distribution that optimizes the worst-case normalized payment with respect to

$$\min_{i \in [n]} \Pr \left(\hat{b}_i \geq b_i \text{ and } \hat{b}_{-i} = b_{-i} \mid b \right) \geq \alpha$$

must take $\Pr(\hat{b} \not\leq b|b) = 0$ and, therefore

$$\min_{i \in [n]} \Pr \left(\hat{b}_i \geq b_i \text{ and } \hat{b}_{-i} = b_{-i} \middle| b \right) = \Pr \left(\hat{b} = b \middle| b \right)$$

implying that it is sufficient to optimize with respect to $\Pr(\hat{b} = b|b) \geq (1 - \gamma)^n = \alpha$.

The following lemmas characterize the welfare and revenue approximations of the reduction generated by a resampling distribution μ :

Lemma D.12 *The welfare approximation of a resampling distribution μ for a bid b is*

$$\alpha = \min_{i \in [n]} \Pr \left(\hat{b}_i \geq b_i \text{ and } \hat{b}_{-i} = b_{-i} \middle| b \right) .$$

Proof: For a bid b , define the set $B^i \subset \mathbb{R}_+^n$ as

$$B^i = \{ \hat{b} | \hat{b}_i \geq b_i \text{ and } \hat{b}_{-i} = b_{-i} \} .$$

Monotonicity of A requires that for all $u \geq b_i$,

$$A_i(u, b_{-i}) \geq A_i(b) .$$

Thus, the allocation received by player i under \mathcal{A} is at least

$$\Pr \left(\hat{b}_i \geq b_i \text{ and } \hat{b}_{-i} = b_{-i} \middle| b \right) A_i(b) = \Pr \left(\hat{b} \in B^i \middle| b \right) A_i(b)$$

and thus the social welfare is at least

$$\begin{aligned} \sum_{i \in [n]} b_i \mathcal{A}_i(b) &\geq \sum_{i \in [n]} b_i \Pr \left(\hat{b} \in B^i \middle| b \right) \mathcal{A}_i(b) \\ &\geq \min_{i \in [n]} \left(\Pr \left(\hat{b} \in B^i \middle| b \right) \right) \sum_{i \in [n]} b_i \mathcal{A}_i(b) . \end{aligned}$$

This lower bound is tight in the following allocation rule

$$A_i(\hat{b}) = \begin{cases} 1 & i = j \text{ and } \hat{b} \in B^j \\ 0 & \text{otherwise} \end{cases}$$

when $j = \operatorname{argmin}_{i \in [n]} b_i \Pr(\hat{b} \in B^i|b)$. ■

Lemma D.13 *The revenue approximation α_R of a reduction given by a resampling distribution μ is bounded from below by the precision*

$$\alpha_P = \inf_b \Pr \left(\hat{b} = b \middle| b \right) \leq \alpha_R$$

and above by

$$\alpha_R \leq \inf_b \min_{i \in [n]} \Pr \left(\hat{b}_i \geq b_i \wedge \hat{b}_{-i} = b_{-i} \middle| b \right) .$$

Proof: To see that the precision $\alpha_P = \inf_b \Pr(\hat{b} = b | b)$ is a lower bound on the revenue approximation, consider decomposing the mechanism produced by the reduction as follows: with probability α_P , the mechanism uses the original allocation function, and with probability $1 - \alpha_P$ it chooses an allocation function A^{rs} that resamples bids more frequently. Since prices are linear, the final expected price will be the weighted sum of the truthful prices for A and the truthful prices for A^{rs} .

For positive types, revenue from both A and A^{rs} will be nonnegative, and the revenue of the resulting mechanism will be the weighted sum of the revenues from A and A^{rs} . Thus, since A is chosen with probability α_P , the revenue of their combination will be at least α_P times the revenue from A .

Next we use the allocation function from Lemma D.12 to give an upper bound. For clarity, we assume that the infimum in the bound of α is attained by some b . (The proof when the infimum is not attained is messier but fundamentally the same.) Let b be a bid such that

$$\min_{i \in [n]} \Pr(\hat{b}_i \geq b_i \wedge \hat{b}_{-i} = b_{-i} | b) = \alpha .$$

Again, let $B^i \subset \mathbb{R}_+^n$ be the set

$$B^i = \{\hat{b} | \hat{b}_i \geq b_i \text{ and } \hat{b}_{-i} = b_{-i}\} ,$$

and consider following allocation function, where $j = \operatorname{argmin}_{i \in [n]} b_i \Pr(\hat{b} \in B^i | b)$:

$$A_i(\hat{b}) = \begin{cases} 1 & i = j \text{ and } \hat{b} \in B^i \\ 0 & \text{otherwise.} \end{cases}$$

When this allocation function is implemented directly with the Archer-Tardos pricing rule, the revenue when bidders say b will be

$$\sum_{i \in [n]} b_i A_i(b) - \int_{-\infty}^{b_i} A_i(u, b_{-i}) du = b_j .$$

Now, for any single call reduction, the expected revenue will be

$$\begin{aligned} \sum_{i \in [n]} b_i \mathbf{E}[A_i^{sc}(b)] - \int_{-\infty}^{b_i} \mathbf{E}[A_i^{sc}(u, b_{-i})] du &\leq b_j \mathbf{E}[A_j^{sc}(b)] \\ &= b_j \Pr(\hat{b} \in B^j | b) . \end{aligned}$$

Thus, the revenue approximation when players bid b is at most $\Pr(\hat{b} \in B^j | b)$. ■

Lemma D.14 *The worst-case bid-normalized payment for a resampling distribution μ is at least*

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \max \left(\frac{1 - \gamma^{(=)}}{\gamma^{(=)}}, \frac{1 - \gamma^{(>)}}{\gamma^{(>)}} \right) \text{ a.e.}$$

where

$$\gamma^{(=)} = 1 - \left(\frac{\Pr(\hat{b} = b | b)}{\Pr(\hat{b} \leq b | b)} \right)^{\frac{1}{n}}$$

and

$$\gamma^{(>)} = 1 - \left(\frac{\min_{i \in [n]} \Pr(\hat{b}_i > b_i \wedge \hat{b}_{-i} = b_{-i} | b)}{\frac{1}{n} \Pr(\hat{b} \not\leq b | b)} \right)^{\frac{1}{n-1}} .$$

The bound holds everywhere under the nice distribution assumption.

Proof: For the sake of clarity, we assume the nice distribution assumption. The general case follows naturally by carrying extra terms through the analysis.

Corollary D.4 says that for any $M \subset [n]$ and $i \notin M$,

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{\pi^\mu(M \cup \{i\}, b)}{\pi^\mu(M, b)}.$$

Since $\sum_{M \subseteq [n]} \pi(M, \bar{b}) = \Pr(\hat{b} \leq \bar{b} | \bar{b})$, applying Lemma D.9 with $\eta(S) = \pi^\mu(M, b)$ implies that

$$\max_{M \subseteq [n]} \frac{\pi^\mu(M \cup \{i\}, b)}{\pi^\mu(M, b)} \geq \frac{1 - \gamma^{(=)}}{\gamma^{(=)}}$$

where $\gamma^{(=)}$ is

$$\gamma^{(=)} = 1 - \left(\frac{\Pr(\hat{b} = b | b)}{\Pr(\hat{b} \leq \bar{b} | \bar{b})} \right)^{\frac{1}{n}}.$$

Thus,

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{1 - \gamma^{(=)}}{\gamma^{(=)}}.$$

Next, define $\nu^\mu(M, j, b)$ as the probability that $\hat{b}_j > b_j$ while bids $i \neq j$ obey M (that is, $\hat{b}_i = b_i$ for $i \in M$ and $\hat{b}_i < b_i$ if $i \notin M$). Lemma D.8 implies that for all j , $M \subseteq [n] \setminus \{j\}$ and $i \notin M \cup \{j\}$,

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{\nu^\mu(M \cup \{i\}, j, b)}{\nu^\mu(M, j, b)} \text{ a.e.}$$

For any particular j , applying Lemma D.9 with $\eta(S) = \nu^\mu(S, j, b)$ as above implies that

$$\max_{M \subset [n] \setminus \{j\}} \frac{\nu^\mu(M \cup \{i\}, j, b)}{\nu^\mu(M, j, b)} \geq \frac{1 - \gamma^{(j)}}{\gamma^{(j)}}$$

where $\gamma^{(j)}$ is

$$\gamma^{(j)} = 1 - \left(\frac{\Pr(\hat{b}_j > b_j \wedge \hat{b}_{-j} = b_{-j} | b)}{\Pr(\hat{b}_j > b_j \wedge \hat{b}_{-j} \leq b_{-j} | b)} \right)^{\frac{1}{n-1}}.$$

Since the probabilities $\Pr(\hat{b}_j > b_j \wedge \hat{b}_{-j} \leq b_{-j} | b)$ are disjoint, there must be some j such that

$$(1 - \gamma^{(j)})^{n-1} \geq \frac{\min_{i \in [n]} \Pr(\hat{b}_i > b_i \wedge \hat{b}_{-i} = b_{-i} | b)}{\frac{1}{n} \Pr(\hat{b} \not\leq b | b)}.$$

$$\frac{\Pr(\hat{b}_j > b_j \wedge \hat{b}_{-j} = b_{-j} | b)}{\Pr(\hat{b}_j > b_j \wedge \hat{b}_{-j} \leq b_{-j} | b)} \geq \frac{\min_{i \in [n]} \Pr(\hat{b}_i > b_i \wedge \hat{b}_{-i} = b_{-i} | b)}{\frac{1}{n} \Pr(\hat{b} \not\leq b | b)}.$$

Thus, it must be that

$$\max_{j, M \subset [n] \setminus \{j\}, i \notin M \cup \{j\}} \frac{\nu^\mu(M \cup \{i\}, j, b)}{\nu^\mu(M, j, b)} \geq \frac{1 - \gamma^{(>)}}{\gamma^{(>)}}$$

where $\gamma^{(>)}$ satisfies

$$\gamma^{(>)} = 1 - \left(\frac{\min_{i \in [n]} \Pr(\hat{b}_i > b_i \wedge \hat{b}_{-i} = b_{-i} | b)}{\frac{1}{n} \Pr(\hat{b} \not\leq b | b)} \right)^{\frac{1}{n-1}}.$$

Consequently,

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{1 - \gamma^{(>)}}{\gamma^{(>)}}$$

as desired. ■

We now have the tools to prove that a resampling distribution that optimizes payments subject to a precision bound also optimizes them subject to a welfare approximation or revenue approximation bound:

Proof:[of Lemma D.3] For clarity, we argue under the nice distribution assumption. Subject to $\min_{i \in [n]} \Pr_\mu(\hat{b}_i > b_i \wedge \hat{b}_{-i} = b_{-i} | b) \geq \alpha > 2^{-n}$, the BKS transformation achieves

$$\sup_{\hat{b}} \left| \frac{\rho_b^{BKS}(\hat{b})}{b_i} \right| = \frac{\alpha^{\frac{1}{n}}}{1 - \alpha^{\frac{1}{n}}},$$

so any optimal distribution must do at least as well.

Let μ be some resampling distribution. If $\Pr_\mu(\hat{b} \not\leq b | b) \neq 0$, either

$$\frac{\Pr_\mu(\hat{b} = b | b)}{\Pr_\mu(\hat{b} \leq \bar{b} | \bar{b})} > \alpha,$$

or

$$\frac{\min_{i \in [n]} \Pr_\mu(\hat{b}_i > b_i \wedge \hat{b}_{-i} = b_{-i} | b)}{\Pr_\mu(\hat{b} \not\leq b | b)} \geq \alpha.$$

In the first case, applying Lemma D.14 gives

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{1 - \gamma^{(=)}}{\gamma^{(=)}} > \frac{\alpha^{\frac{1}{n}}}{1 - \alpha^{\frac{1}{n}}} = \sup_{\hat{b}} \left| \frac{\rho_b^{BKS}(\hat{b})}{b_i} \right|$$

and therefore μ cannot be optimal.

In the second case, Lemma D.14 and the assumption that $\alpha > 2^{-n} \geq \frac{1}{n^n}$ gives

$$\begin{aligned} \gamma^{(>)} &\leq 1 - (n\alpha)^{\frac{1}{n-1}} \\ &< 1 - (\alpha^{-\frac{1}{n}} \alpha)^{\frac{1}{n-1}} \\ &= 1 - \alpha^{\frac{1}{n}}. \end{aligned}$$

Thus, $\gamma^{(>)} < 1 - \alpha^{\frac{1}{n}}$, so

$$\sup_{\hat{b}} \left| \frac{\rho_b^\mu(\hat{b})}{b_i} \right| \geq \frac{1 - \gamma^{(>)}}{\gamma^{(>)}} > \frac{\alpha^{\frac{1}{n}}}{1 - \alpha^{\frac{1}{n}}} = \sup_{\hat{b}} \left| \frac{\rho_b^{BKS}(\hat{b})}{b_i} \right|$$

so again μ cannot be optimal.

It follows that any optimal distribution μ must have $\Pr(\hat{b} \not\leq b|b) = 0$ and, therefore

$$\min_{i \in [n]} \Pr \left(\hat{b}_i > b_i \wedge \hat{b}_{-i} = b_{-i} \mid b \right) = \Pr \left(\hat{b} = b \mid b \right) .$$

Thus, a distribution which wishes to optimize the worst-case normalized payment subject to $\Pr(\hat{b} = b|b) \geq \alpha$ will also optimize payments subject to $\min_{i \in [n]} \Pr(\hat{b}_i > b_i \wedge \hat{b}_{-i} = b_{-i}|b) \geq \alpha$, and will have $\Pr(\hat{b} = b|b) = \min_{i \in [n]} \Pr(\hat{b}_i > b_i \wedge \hat{b}_{-i} = b_{-i}|b)$. ■

E Analysis Definitions, Facts, and Lemmas

This section provides a limited background on analysis concepts.

E.1 Measures and Integrals

We begin with various possible set of axioms a collection of sets may satisfy, and their technical names.

Definition 9 (σ -algebra) *The σ -algebra over a set U is a non-empty collection Σ of subsets of U that is closed under complementation and countable union of its members. The pair (U, Σ) is called a measurable space.*

Definition 10 (Generated σ -algebra) *Given a set U and a collection of subsets F of U , there is a unique smallest σ -algebra over U containing all the elements of F . This σ -algebra is denoted by $\sigma(F)$ and is called as the σ -algebra generated by F .*

Definition 11 (Borel σ -algebra) *The Borel σ -algebra $\mathcal{B}(U)$ of a metric space U is the σ -algebra generated by the collection of all open sets of U .*

Definition 12 (Measurable sets) *Once we fix a measurable space (U, Σ) , the sets $X \in \Sigma$ are called measurable sets.*

Definition 13 (Measurable functions) *Given two measurable spaces (U, Σ) and (U', Σ') , a function $f : U \rightarrow U'$ is measurable if for each $X' \in \Sigma'$, $f^{-1}(X') \in \Sigma$.*

We are now ready for the definition of a measure.

Definition 14 (Measure) *Given a measurable space (U, Σ) , we equip it with a measure ν , which is function $\nu : \Sigma \rightarrow [0, \infty]$ that satisfies*

1. $\nu(\emptyset) = 0$
2. *Countable additivity, i.e. for all countable sequences $\{X_i\}_{i \in \mathbb{Z}}$ of pairwise-disjoint sets in Σ , $\nu(\cup_{i \in \mathbb{Z}} X_i) = \sum_{i \in \mathbb{Z}} \nu(X_i)$.*

A measure ν is said to be finite if $\nu(U)$ is finite.

Definition 15 (Probability measure) *A measure is a probability measure if $\nu(U) = 1$.*

Definition 16 (Signed measure) *A signed measure is a function $\nu : \Sigma \rightarrow [-\infty, \infty]$ that satisfies $\nu(\emptyset) = 0$ and countable additivity.*

Fact E.1 *If ν_1 and ν_2 are finite (signed) measures, then $\nu_3(X) = \nu_1(X) - \nu_2(X)$ is a finite signed measure.*

Convention According to standard convention, a measure is not signed unless explicitly stated. For the purposes of this paper, the set U will always be \mathbb{R}^n .

Apart from the set collections defined via σ -algebras, we also need some weaker set collections, which we define below.

Definition 17 (π -system) The π -system over a set U is a non-empty collection P of subsets of U that is closed under finite intersection of its members, i.e., $X_1 \cap X_2 \in P$ whenever X_1 and $X_2 \in P$.

Definition 18 (λ -system, or Dynkin system) The λ -system over a set U is a non-empty collection L of subsets of U that is closed under complementation and countable disjoint union of its members.

Fact E.2 (Dynkin's theorem) If P is a π -system and L is a λ -system over the same set U , and $P \subseteq L$, then $\sigma(P) \subseteq L$, i.e., the σ -algebra generated by P is contained in L .

The Hahn and Jordan decompositions decompose a signed measure into two measures. They will be useful when we discuss the integral with respect to a signed measure.

Fact E.3 (Hahn decomposition theorem) The Hahn decomposition of a signed measure ν over a measurable space (U, Σ) consists of two sets $P, N \in \Sigma$ such that $P \cup N = U$, $P \cap N = \emptyset$, and for all measurable sets $X \subseteq P$, $\nu(X) \geq 0$ and for all measurable sets $X \subseteq N$, $\nu(X) \leq 0$. The Hahn decomposition is guaranteed to exist and be unique (up to a set of measure 0)

Fact E.4 (Jordan decomposition theorem) This theorem is a consequence of Hahn decomposition theorem, and states that every signed measure ν can be decomposed as two (non-negative) measures $\nu^+(X) = \nu(X \cap P)$ and $\nu^-(X) = -\nu(X \cap N)$, where P and N are the Hahn decomposition of ν . The measures satisfy $\nu(X) = \nu^+(X) - \nu^-(X)$. The Jordan decomposition is guaranteed to exist and to be unique, and at least one of ν^+ and ν^- is guaranteed to be a finite measure. If ν is finite, then both ν^+ and ν^- are finite.

Definition 19 (Characteristic Function) The characteristic function $1_S(x)$ of a set S is the function that is 1 if $x \in S$ and zero elsewhere, i.e.

$$1_S(x) = \begin{cases} 1, & x \in S \\ 0, & \text{otherwise.} \end{cases}$$

Definition 20 (Simple Function) Given a measurable space (U, Σ) , a function $s : U \rightarrow \mathbb{R}$ is a simple function if it can be written as a finite linear combination of indicator function of measurable sets. That is,

$$s(x) = \sum_{k=1}^n a_k 1_{S_k}(x)$$

for finite sequences of measurable sets $\{S_k\} \in \Sigma$ and coefficients $\{a_k\} \in \mathbb{R}$.

Fact E.5 For any non-negative, measurable function f , there is a monotonic increasing sequence of non-negative simple functions $\{s_k\}$ such that

$$f = \lim_{k \rightarrow \infty} s_k .$$

Definition 21 (Integral) Given a measurable space (U, Σ) , the integral of a function $f : U \rightarrow \mathbb{R}$ with respect to a measure ν is defined incrementally. For any measurable set X , the integral of 1_X is

$$\int_U 1_X d\nu = \nu(X) .$$

For any simple function $s : U \rightarrow \mathbb{R}$,

$$\int_U s d\nu = \sum_{k=1}^n a_k \nu(X_k) .$$

For a general non-negative function $f : U \rightarrow \mathbb{R}$,

$$\int_U f d\nu = \sup \left\{ \int_U s d\nu \quad : \quad 0 \leq s \leq f \text{ and } s \text{ is simple} \right\} .$$

For general f , let $f^+(x) = \max(f(x), 0)$ and $f^-(x) = \max(-f(x), 0)$, i.e. f^+ and f^- are the positive and negative parts of f respectively. Then

$$\int_U f d\nu = \int_U f^+ d\nu - \int_U f^- d\nu .$$

Finally, for some measurable set Y ,

$$\int_Y f d\nu = \int_U f d\nu_Y$$

where $\nu_Y(X) = \nu(U \cap Y)$.

Fact E.6 (Monotone Convergence Theorem) For any countable, monotone sequence of measurable functions $\{f_k\}$ (that is, sequence where $f_k \geq f_{k-1}$ pointwise),

$$\lim_{k \rightarrow \infty} \int f_k d\nu = \int \lim_{k \rightarrow \infty} f_k d\nu .$$

The following fact follows because $g_k = \sum_{i=1}^k f_i$ satisfies the monotone convergence theorem:

Fact E.7 For any countable sequence of nonnegative measurable functions $\{f_k\}$

$$\sum_{k=1}^{\infty} \int f_k d\nu = \int \sum_{k=1}^{\infty} f_k d\nu .$$

Fact E.8 Let $\{X_k\}$ be a countable sequence of disjoint sets. Then

$$\sum_k \int_{X_k} f d\nu = \int_{\cup_k X_k} f d\nu .$$

Definition 22 (Integral with respect to a Signed Measure) The integral of a function f with respect to a signed measure ν is

$$\int_U f d\nu = \int_U f d\nu^+ - \int_U f d\nu^- ,$$

where ν^+ and ν^- are the Jordan decomposition of ν .

E.1.1 Densities and Derivatives

Definition 23 (Absolute continuity) Given a signed measure ν and a measure μ on the same measurable space, ν is absolutely continuous w.r.t. μ , if for every measurable set V where $\mu(V) = 0$, we have $\nu(V) = 0$.

We now state below the Radon-Nikodym theorem the way we use it, though the theorem itself is more general.

Fact E.9 (Radon-Nikodym Theorem) The Radon-Nikodym theorem states that given a finite signed measure ν and a finite measure μ on the same measurable space such that ν is absolutely continuous w.r.t. μ , the measure ν has a density, or “Radon-Nikodym derivative”, with respect to μ , i.e., there exists a μ -measurable function ρ taking values in $[0, \infty]$, such that for any μ -measurable set X we have

$$\nu(X) = \int_X \rho d\mu .$$

Fact E.10 If ρ is a Radon-Nikodym derivative of measure ν w.r.t. measure μ , then

$$\int_X f(x) d\nu = \int_X \rho(x) f(x) d\mu$$

wherever $\int_X f(x) d\nu$ is well defined.

E.2 Almost Everywhere

Definition 24 (Almost Everywhere) A property $P(s)$ is said to hold almost everywhere on a set S if the subset of S on which $P(s)$ is false has measure zero (or is contained in a set that has measure 0). It is abbreviated a.e.. The exact measure used will become clear from the context.

Definition 25 (Almost Surely) If a property $P(s)$ is false with probability 0 with respect to some distribution over s , then it is said to hold almost surely. This is equivalent to saying $P(s)$ is true almost everywhere with respect to the probability measure associated with the distribution.

Fact E.11 For a non-negative measurable function f and measure μ , $\int f d\mu = 0$ if and only if $f(x) = 0$ almost everywhere.

Fact E.12 For any measurable function f and signed measure ν , if $\int_X f d\nu = 0$ for all measurable X , then $f = 0$ almost everywhere.

The second fact follows from the first by a standard argument — decompose f into its positive and negative parts and decompose ν according to its Hahn decomposition. This partitions the space into four sets over which the integral may be written as a non-negative function with respect to a non-negative measure. Apply Fact E.11 to each of the four sets.

E.3 Extrema

Definition 26 (Supremum/Infimum) For a set S , the supremum of S , denoted $\sup S$, is the smallest value x such that $x \geq s$ for all $s \in S$. Similarly, the infimum of s is the largest value x such that $x \leq s$ for all $s \in S$.

Definition 27 (Limit Superior/Inferior) For a real-valued function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the limit superior, denoted $\limsup_{u \rightarrow b} f(u)$, may be defined as follows:

$$\limsup_{u \rightarrow b} f(u) = \lim_{\epsilon \rightarrow 0} \left(\sup_{u \in BALL(b, \epsilon)} f(u) \right)$$

where $BALL(b, \epsilon)$ is the open ball of radius ϵ centered at b . It is an upper bound on the limit of $f(u_i)$ for any sequence of values $\{u_i\}$ that converges to b . The \liminf is defined similarly. Note that while the limit may not exist as $u \rightarrow b$, the \limsup and \liminf are always well defined for real-valued functions.

It is natural to generalize \sup and \limsup to almost everywhere:

Definition 28 (Essential Supremum/Infimum) The essential supremum of a set S , denoted $\text{ess sup } S$, is the smallest value x such that $x \geq s$ almost everywhere, i.e. the set of values $T = \{s | s \in S \text{ and } s > x\}$ has measure zero. The essential infimum ess inf is defined similarly.

Definition 29 (limesssup/limessinf) For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the $\lim \text{ess sup}_{u \rightarrow b} f(u)$ can be defined as follows:

$$\lim \text{ess sup}_{u \rightarrow b} f(u) = \lim_{\epsilon \rightarrow 0} \left(\text{ess sup}_{u \in BALL(b, \epsilon)} f(u) \right) .$$

It can be understood as a version of the \limsup that will ignore values that $f(x)$ only attains on sets with measure zero. The $\lim \text{ess inf}$ is defined similarly. Like the \limsup and \liminf , the $\lim \text{ess sup}$ and $\lim \text{ess inf}$ are always well defined for real-valued functions.